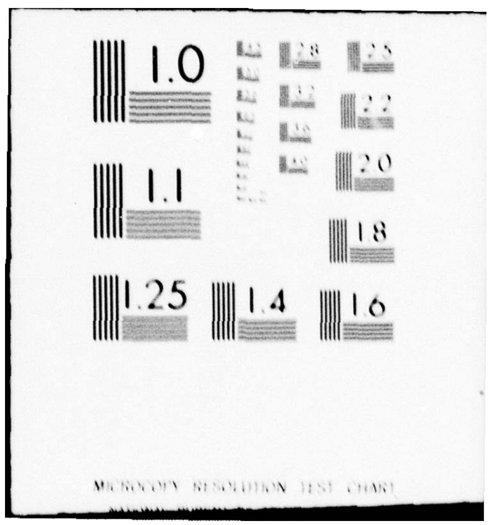
AD-A057 835 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 5/9
METHODS TO ASSESS WORKLOAD.(U)
JUN 78
UNCLASSIFIED AGARD-CP-216 NL

1 OF 2
AD A057 835


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
MICROCOPY RESOLUTION TEST CHART

AGARD-CP-216

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No. 216

Methods to Assess Workload

NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION AND AVAILABILITY
ON BACK COVER

1

14 AGARD-CP-216

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

11 Jun 78

AGARD Conference Proceedings No.216

6 METHODS TO ASSESS WORKLOAD.

9 Conference proceedings,

12 138p.

Papers presented at the Aerospace Medical Panel Specialists' Meeting
held in Cologne, FRG, 18–22 April 1977.

400 043

78 08 22 049

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Program and the Aerospace Applications Studies Program. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced directly from material supplied by AGARD or the authors.

Published June 1978

ISBN 92-835-1285-5

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London, W1P 1HD

PREFACE

With the evolution of advanced aircraft containing a vast array of semi-automated, computerized subsystems, displays, and functions, and the emergence of multi-mission concepts and roles, pilot workload has become of increasing concern. It is a focus of attention from initial design to full scale employment. The workload issue confronts the design engineers, the production engineers, and the test and evaluation engineers during development. It confronts commanders at all operational levels when an aircraft comes into service. But this is not the end of the problem. Workload issues periodically reappear throughout the service life of an aircraft, as modifications and upgrades are made, as the aircraft is moved into new operational environments, and as new roles and missions evolve.

Each of the groups referenced above has its own techniques for assessing and resolving workload problems. These techniques have demonstrated value. However, the specialists in the aerospace medical sciences are primary consultants and are repeatedly enlisted for the analysis and solution of workload issues, since the focus is on the man in the system, and "man" is our specialty.

The measurement of workload poses many problems. A wide variety of methods are employed, frequently in an interdisciplinary setting. This symposium was organized around methodology questions in order to narrow it down to a manageable segment. Even so, the range of topics is clearly going to be extensive: (a) the study environment, ranging from the laboratory to simulator and inflight studies; (b) the instrumentation and associated techniques; (c) global issues like measurement sensitivity, reliability, and validity. The list is long and the task presented to the participants in the symposium is indeed large.

It should be noted that this was one of two symposia on pilot workload conducted back-to-back during this specialists' meeting. This one focused on methods; the other, chaired by Auffret (France), brought to gether the most recent studies in the AGARD aeromedical community. There is, unavoidably, an overlap in the content of the two symposia--an author in the methodology symposium can best present a method to measure workload by reporting a study in which he applied the method; conversely, another author in the "current studies" symposium will report on his methods in the course of describing his study. It is recommended, therefore, that the reader treat these two conference proceedings (CP 216 on methods and CP 217 on studies) as companion pieces on the status of workload research at this time.

AEROSPACE MEDICAL PANEL

Panel Chairman:	Médecin Général G.Perdriel, FAF
Panel Deputy Chairman:	Air Commodore J.N.C.Cooke, RAF
Panel Executive:	Lt Colonel F.Monesi, IAF

MEETING ORGANIZATION

Host Coordinator and Program Organizer:	Dr med. K.E.Klein

SESSION ORGANIZER

Dr B.O.Hartman, US

TECHNICAL EVALUATION REPORT (TER)

The focus of this symposium is on methods of measuring workload. The focus of the companion symposium is on studies--data collection and findings. However, the distinction is not that clear. Studies necessarily reveal the methods employed. Descriptions of method frequently include data, to further amplify the method and its utility. It seems desirable, therefore, to include some material from the other symposium in this TER.

Given the emphasis on methods, it seems appropriate to structure this technical evaluation in a methodical way. I suggest that methodology can be approached with a matrix. Many matrices are possible. I propose one in which one axis be the arena in which data are collected and the other be the nature of the measures. The level of detail on each axis is, of course, a problem; for the sake of simplicity I propose broad categories, leaving it to others to develop greater detail as needed.

The categories chosen for the axis describing the arena in which the data are collected are proposed as follows:

(a) in a model

(b) in a laboratory

(c) in a simulator

(d) in the field

(e) in flight

The continuum is intended to transition from the least to the most realistic (model to in flight).

The categories chosen for the nature of the measures are proposed as follows:

(1) measures of system performance

(2) measures of pilot performance

(3) analogues of pilot performance

(4) measures of pilot status

The nature of the continuum underlying these categories is more difficult to describe; I suggest that it proceeds from the category of most immediate interest to the operational NATO community we support to that of least immediate interest (system performance to pilot status). It should be said that if I specified an audience different from the operational community (e.g., chose the AGARD scientific community), this continuum would undoubtedly look much different.

A matrix has been created. As the next step, each paper from the two symposia will be entered into the appropriate cell. I will be arbitrary, and enter a paper only once, again for simplicity.

WORKLOAD MEASUREMENT METHODOLOGY MATRIX

	1. Measures of system performance	2. Measures of pilot performance	3. Analogues of pilot performance	4. Measures of pilot status	Examples*
A. in a model	A-13	B-10		(B-1)	ton-miles kill ratios attrition
B. in a laboratory			A-4	A-8 A-10	reaction time tracking scores perceptual effic.
C. in a simulator		A-11 A-12 B-9			procedural error emergency response glide path deviation
D. in the field			B-11 B-12	A-6 A-7 B-7 B-8	sortie rate in-commission rate cargo pass-thru time
E. in flight	A-2 B-4	A-1 A-3 B-2 B-3	B-6	A-5 A-9 B-5	flight path deviations eye movement patterns crew activity

Examples*	altitude control	control movements	pilot opinion	secondary tasks
	navigation	visual scanning	synthetic tasks	neurophys. status
	gunnery scores	communication	traditional tasks	biochem. status

*generic--not from this symposium

Three problems were revealed as this categorization was performed:

(1) The repeated temptation to enter a paper in more than one cell despite this ground rule: each paper is to be entered only once.

(2) Some difficulty in deciding which cell was most appropriate for a specific paper.

(3) Rarely, examining a paper for which there seemed to be no reasonable fit in the matrix. Paper B-1 (cell A-4) is enclosed in parentheses because of this problem.

These suggest that the matrix and the way it is used needs further development.

Overall, however, this matrix analysis reveals a substantial spread of research activity across the matrix. This would have been even more evident had multiple entries per paper been allowed. It is the opinion of this evaluator that this is healthy, because it allows balanced service to the operational community (whose interests lie most strongly in the lower left corner), the general AGARD scientific community (whose interests lie largely in the central portions of the matrix), and the AGARD medical community (whose interests lie mostly in the right column). The balance is good.

CONTENTS

IS MAN THE WEAKEST LINK?

By

Harvey G. Gregoire, Ph.D.
Lieutenant Commander, MSC, U. S. Navy
Naval Aerospace Medical Research Laboratory
Pensacola, Florida 32508

A portable video tape recorder and camera were installed on the aft bulkhead of the crew compartment to obtain inflight video records of aircrew activity during test flights of a new antisubmarine warfare (ASW) aircraft. The technique proved successful in providing graphic records of real time crew utilization of systems controls and displays. In combination with crew briefing and debriefing interviews, the annotated tapes provided a basis for activity sampling, scoring task times and error rates, evaluating display utilization, counting control inputs and display mode changes, and for other observations pertinent to crew task loading. Data were obtained during passive acoustic submarine search and track operations during actual ASW test flights. These data indicated that during passive acoustic search and tracking, individual control tasks requiring one or more pushbutton operations were performed from 6 to 10 times per minute. Display mode changes occurred at intervals of approximately two minutes. The operators in this limited sample made Integrated Control System inputs exclusively with the right hand. With the exception of a few particular pushbutton inputs, every individual pushbutton operation was guided visually. Keying tasks which occurred with the highest frequency were KEYBOARD numerical entries. Based on the limited operator sample and the video tape and interview procedure, the following potential applications are identified.

1. Capability of real time activity recording with slow time analysis.

2. Actual in-flight data sampling.

3. Capability for accurate and detailed time-line analysis to be correlated with human engineering design limitations of controls and displays.

The advantages of small battery-powered video tape analysis during in-flight operations are presented with emphasis on workload relationship to human engineering. Of particular interest are the differences between aircrew debrief accuracy with vs without the videotape utilization. Specific results from actual test flights using this procedure are discussed.

The speed of technological advancement in the development of aircraft systems designed to locate and destroy enemy targets is unparalleled today. In recent years, the increased sophistication of airborne computers, avionics, and sensing systems has progressed from idea to blueprint to hardware more rapidly than thought possible only a few short years ago. Engineers are progressively developing avionics weapon systems that provide airborne operators huge quantities of relevant functional data necessary to fulfill search, location and destroy mission objectives. However, is it possible that man, the developer of unprecedented capabilities found in sensors, computers, and automated systems, has provided more information than man the operator can process in some instances? Is there such an emphasis on hardware system information processing and presentation that managers and designers have forgotten to consider <u>human</u> system limitations?

As the complexity of the man-machine interface evolves, these questions must <u>not</u> go unanswered, if we are to prevent mission degradation.

Limited to an analysis of an aircraft weapon system, this paper will describe an approach used to identify human system overloading, point out where human processing is overburdened and suggest possible design alternatives that will ameliorate weapon system efficacy. Due to the requirements of keeping this paper unclassified, methodology will be discussed without mention of any aircraft system specific information. Thus, crewman or system functions will simply be called "decision A, system B", etc.

The aircraft system was a modern ASW (Antisubmarine Warfare) jet carrier-based aircraft designed to locate, identify, and, if necessary, destroy enemy submarines in a task force environment. In general, sonobuoy sensors are dropped from the aircraft and data are relayed back to the aircraft. These data are analyzed and processed to

---

*Opinions or conclusions contained in this report are those of the author and do not necessarily reflect the views or endorsement of the Navy Department.

permit target identification, localization and destruction. The TACCO or Tactical Coordinator is a Naval Flight Officer who functions as the primary decision maker in this airborne weapon system. The TACCO is the operator responsible for applying the full capabilities of his airborne platform and crew against a sophisticated, silent, and speedy adversary capable of evasion throughout the expanses and depth of oceans and seas. For the purposes of this paper, the TACCO will hereafter be referred to as the "operator."

To assist in identifying problems of operator overload during times of increased mission activity, it is essential to accurately and comprehensively describe in task analysis form the workload elements which transpire during general and specific subtasks that combine to make up the massive quantity of observations, responses, decisions, evaluations, and actions accomplished by aircrew operators during mission performance.

It should be noted that the reason for developing a task analysis methodology in this case was twofold. First, a requirement existed to determine if the physical arrangement of numerous controls and displays in the subject aircraft was optimally designed to facilitate operator interface during peak workloads. This requirement is a human engineering sub-element to be evaluated during preliminary flight testing prior to fleet acceptance of aircraft. Secondly, the organization tasked with this workload analysis had unfortunate prior experience in this area of testing during a previous weapons system evolution. The prior experience consisted of subcontracting the workload analysis to a separate organization. The prior effort also concerned itself with the TACCO operator and a sophisticated set of computer-assisted avionics and information processing equipments to detect and destroy enemy submarines. Although the prior effort involved a larger, different aircraft (turbo-prop vs jet), the workload analysis problems were similar. The prior subcontracted investigators boarded the aircraft during evaluation flights flown against live (friendly) submarines. Armed with a multitude of stopwatches, forms, and clipboards, their intent was to manually record, time, and annotate all the operator-crew-equipment interactions during the contact, localization, and attack portions of the mission using classical time-line techniques. The results were disastrous. To paraphrase the investigator's own comments: "Ten minutes into the exercise we were a half hour behind in our attempts to record what was happening." It became obvious that stopwatches, recording forms and pencils were not the tools to keep pace with the tempo of live ASW operations. Actions, movements, and decisions, which could be stopped and analyzed in a laboratory setting, could not be interrupted in the sky during live missions. With all due respect to simulation--what happens in the sky during live missions is the relevant issue in task analysis.

With the history of the problems associated with the previous attempts, a methodology was planned to sample activity during live flights. The idea was not unique, original, or exotic. The investigator, recalling instant replays on televised football games, decided such an approach might be the answer in how to <u>accurately</u> record activity during actual operations and then (after the fact) interrupt such activity in order to dissect and analyze any subtask or control input and associated decisions made by the operator. The notion of using a sound-recording motion picture film camera was discarded due to the size, weight, reloading, and film processing requirements involved. Although the investigator was sure that the videotape approach was not a unique idea, it was no easy task to get approval for installation of a small battery-powered TV camera in the prototype aircraft to be tested since no precedent could be found for this specific application. Subsequent to design and fabrication of camera mounts and devices to secure camera power sources, the system was tested under static ground conditions. It was discovered that the ambient light at the operator station had a significant effect on the quality of the video picture obtained. After much trial and error in partial shielding of dome lights with cardboard, filters over portholes, etc., an excellent ambient light balance was achieved between operator requirements and good quality video pictures. Inflight videotapes were then made of the operator during operations against actual submarines on

The technique proved successful in obtaining real time, uninterrupted visual and audio records of operator activity during actual operations. With the replay of the videotape occurring simultaneously with a post mission debrief, an expanded, overlay tape recording could be made which could explain all the activity, decision-making, etc., that went into an element as simple and quick as pressing a pushbutton. For example, a videotape might portray the operator pausing 4.3 seconds between pressing two separate buttons to display certain data on his tactical display. The rationale of why the specific second control was activiated in preference to some other control may take several minutes of verbal explanation or many pages of written text. The point is that once the operator's live mission activity is permanently recorded, any mission segment no matter how large or small can easily be replayed over and over again to analyze carefully and detect small details that the human eye may not accurately perceive nor the human memory recall during busy and hurried moments of operator activity.

This point was vividly illustrated in one instance during initial application of operator workload video recording. The investigator reviewed a mission videotape prior to an operator debrief which was to be combined with a second video replay. Prior to the debrief, the investigator remarked to the operator (who was highly experienced) that it appeared that extremely limited use of the left hand for control of data inputs was observed in the C, D, and E phases of the mission (high workload times). The operator assured the investigator that this was not so since he "remembered" certain pushbutton control inputs that he was "sure" he had made with his left hand. Upon reviewing the videotape during the workload debrief, the operator repeatedly shook his head in disbelief when he realized that, in fact, <u>never</u> <u>once</u> during the subject C, D, and E phases of the mission did he <u>ever</u> use his left hand. This occurred less than an hour since landing from the mission in question

Lesson #1: During periods of sustained intense activity operators cannot recall actions or thought processes accurately enough to facilitate a workload analysis.

Identification of this disproportionate keyboard function loading for one hand led to relocation of several functions (pushbuttons) to the left side of the keyboard to share limb loading during busy mission phases.

Additionally, it was noted several times with various observers that in one particular type of tactical situation, particular incorrect control data inputs were made as a result of displayed information that had been misinterpreted. Upon reviewing videotapes the second or third time, the operators concluded that the information displayed from system X was correct, but the format of information display caused erroneous interpretation.

Lesson #2: During periods of sustained intense activity discrimination errors increase.

Or, to state it in other words, busy operators will believe what they think they saw. When seeing the same data again in a situation where time is not critical and can be "stopped" or replayed if necessary, better discrimination and decision-making can take place. In the circumstance just described, a display format change was made to solve the problem.

Lesson #3: By being able to "relive" the rapid airborne circumstance via repeated, unhurried small portions of video playback, relatively undetectable discrimination and decision errors can be identified.

Operator-machine interface in this situation was primarily accomplished through alpha-numerical keyboard entries. Because alphanumeric keys were not distinguishable from other keys by space, shape, or color coding, the operators were required to direct visual attention away from other displays to look at each alphanumeric key before making an entry. Most operators remarked that during operations they had not been conscious of the time lost or disruptive effect of refocusing visual (and mental) attention to the numerical keys. However, the frequent small distractions caused by this requirement were noticeable to the operators when they subsequently saw themselves operate on video replay.

Lesson #4: During periods of sustained intense activity, minor distractions can increase discrimination errors. Such distractions may not be initially detectable because of the very reason that they occur in a context of intense sustained activity as opposed to a more normal pace of activity.

Additional videotape analysis revealed several other display/control relationships requiring unnecessarily cumbersome sequences of keyboard inputs which were correctable via software changes.

Lesson #5: Only in the activity sequences and time compressions of actual operations can particular task procedures and equipment/operator relationships be identified as non-optimum.

For the typical workloads evaluated by video sampling during these flight tests, the value of correct operator task performance to overall system tactical performance was not seriously degraded by correction of those deficiencies described above.

A brief summary of the lessons derived from applying video sampling to operator task analysis includes the following observations:

1. Human memory is not reliable when recalling certain details from periods of heavy workload.

2. Visual discrimination (and probably other modes) is degraded during high rates of activity.

3. Repeating (replaying) portions of high activity rate events at slower speeds facilitates error detection which might otherwise go unnoticed.

4. Distractions not perceptible in some workload environments may be obvious when seen again in a non-hurried (video replay) atmosphere.

5. Only in sampling real time live operations may all the operator/equipment work relationships and interactions be meaningfully analyzed.

To terminate these remarks and to recall the original question of whether or not man is the weakest link in our modern weapons systems, the following conclusions were reached as a result of applying the video task-analysis methodology.

In the case cited, man was not the weakest link. However, the diversity and complexity of data which are acquired, interpreted, and processed by crewmen in modern weapons systems must be optimized by meticulous human engineering of controls and displays. Task loadings which result from inadequate operator-equipment interfaces do have the potential of overburdening man regardless of weapons systems technological sophistication. Improvements in the ability to identify and quantify human workload under live operational circumstances may be our last opportunity to prevent man from becoming the weakest link in future weapons systems.

METHODOLOGICAL CONSIDERATIONS OF VISUAL WORKLOADS OF HELICOPTER PILOTS

by

Ronald R. Simmons and Kent A. Kimball
US Army Aeromedical Research Laboratory
Aviation Psychology Division
Fort Rucker, Alabama 36362

SUMMARY

This report was initiated to review the techniques and modifications developed by the U.S. Army Aeromedical Research Laboratory for assessing visual performance/workload of pilots during helicopter operations. Although the corneal reflection technique for gathering eye movement data is not new, innovative modifications had to be developed to permit accurate data collection in this flight environment. This study reports on these techniques, modifications, and applications.

INTRODUCTION

Evolving from the Army's modern air mobility concept, the helicopter has become a strategic element of the tactical structure. The helicopter is no longer utilized exclusively for air transportation as its fixed wing predecessors were, but is presently also a platform for armament, a vehicle for reconnaissance, and an unsurpassed mode for rapid evacuation of wounded personnel from the combat environment. As the mission of this vehicle becomes more complex, so do the tasks of the pilots who fly them. Pilots no longer simply manipulate the aircraft, but now share this duty with other responsibilities dictated by mission requirements. Little is known about aviator visual and motor workload during helicopter flight under varying mission profiles, and even less can be predicted about the available free time which a pilot can utilize for secondary mission tasks.

The pilot's ability to manipulate his aircraft in a tactical setting is directly related to the inputs or cues he receives from the flight environment. Of those perceptual inputs required to fly the aircraft, visual cues are considered vital. Processing and integrating these cues allow the pilot to detect the aircraft's relative stability, ground reference, and response to his control inputs. During flights conducted under instrument meteorological conditions (IMC), the lack of cues from the environment outside the aircraft requires the pilot to obtain the necessary visual information from instrument displays. As a consequence, there exists the need, independent of visual conditions, to determine what cues are required to achieve maximum pilot efficiency for safe mission accomplishment.

A great variety of apparatus and techniques have been developed for the study of visual performance/workload.[1,2,3] However, the state-of-the-art in eye movement recording instrumentation is still in its infancy. One of the earliest devices was a smoked-drum Kymograph attached to the sclera of the eyeball via fine wire and barbed hooks. During the 1930's the electro-oculography (EOG) technique was developed which utilized electrodes placed around the eyes on the facial structure to monitor differential voltages as the eyeball was rotated.[4]

The earliest documented technique for measuring the visual performance of pilots was to simply record pictures of the human operator's face while he scanned the instruments.[5] Improvements of this method were accomplished by arranging mirrors on the instrument panel and photographing the total arrangement. Documentation of eye movement was obtained by means of a camera mounted behind the pilot. During analysis, a photo interpreter scanned the film to determine which mirror reflected the eye of the pilot at various times during the flight.[6]

This technique was further refined by Mackworth.[7] His approach was to mount a light weight moving picture camera beside the operator's head along with a series of mirrors which reflected a dot representing the eye's motion. This dot was superimposed on photographs of the scene directly in front of the head's centerline. More recently this same "corneal reflection" technique has been utilized by the U. S. Army Aeromedical Research Laboratory in the study of Army pilot visual performance during helicopter flight.[8,9]

Because of the smooth spherical front surface of the cornea, an incident beam of light can be partially reflected forming a bright spot or "highlight" on the cornea. The angle of the reflected light depends upon the angle between the incident light ray and a plane tangent to the reflecting surface. Since the cornea forms an eccentric bulge on the nearly spherical eyeball, the angle of this tangential plane on the cornea at any one point changes as the eye rotates about its center during eye movement. As a result, the position of the highlight follows the direction of movement of the cornea. The reflected beam is easily photographed on film. By mounting a camera lens on a subject's head slightly above and between his eyes, the subject's normal visual field can be recorded, and the highlight can be superimposed on the scene to give a constant eye reference in the subject's field of view. By recording both the visual field and the eye's highlight, the areas of visual concentration and the percent of time for eye stabilization during any flight maneuver can be recorded.

Past research has demonstrated two major advantages of the corneal reflection technique for studying eye movement. First, the method is convenient for large scale testing of subjects in that it requires minimal training. Second, these studies have reported no significant interference with normal eye movement.[10,11]

The purpose of this report is to provide a description of the modifications which have been made to the Eye Mark recording apparatus to improve its data acquisition efficiency for helicopter flight and to delineate the methodology which has been developed for the field application of the device subsequent to

these modifications. This method has been instrumental in obtaining baseline information on pilot visual workload during various helicopter operations, including instrument flight (IFR), visual flight (VFR), and terrain flight. The application of this information to the development of more efficient training techniques, procedures and aircraft instrumentation will provide a significant reduction in the overall visual workload of the aviator during helicopter operations.

METHOD

Apparatus. The equipment utilized to record visual performance by the corneal reflection technique included the NAC Eye Mark Recorder, a LOCAM high speed motion picture camera, special high speed film, and the Helicopter In-flight Monitoring System (HIMS).

NAC Eye Mark Recorder. The basic device employed to study visual performance/workload was the NAC Eye Mark Recorder. By utilizing the NAC, the viewing point and movement of the eye can be detected and recorded. Through this optical device an illuminated reticle is focused on the cornea and reflected by mirrors to record movement of the eyeball such that the reticle always coincides with the eye viewing point. This illuminated reticle is superimposed on a primary image and may be recorded on 16mm film. The general specifications of this system are presented in Table 1 and an illustration of the NAC is provided in Figure 1.

TABLE 1

General Specifications of the NAC Eye Mark Recorder

Field-of-View:	30^{0} type; vertical 22.7^{0}, horizontal 31.4^{0} 60^{0} type; vertical 43.5^{0}, horizontal 60^{0} (Head can be freely moved.)
Eye Mark Size:	0.5mm (.02 in.) V mark on 16mm film
Aperture:	T8
Distance Between Subj & Object:	25cm (9.84") - infinity
Eye Mark Correction Range:	Full frame of 16mm film
Spot Lamp (Eye lamp):	Tungsten lamp Life 100 hours (at 3V use)
Mounting Adjustment:	The device can be adjusted to fit to any head size.
Distance from Eyeball to Half Mirror:	30^{0} type; 35mm - 45mm (1.38" to 1.78") 60^{0} type; 45mm - 55mm (1.78" to 2.16")
Parallax Adjustment Range:	15^{0} downward
Optical Fiber:	Single strand diameter 20mm Effective picture 4 x 5mm Total length 1000mm (39.37")
Power Supply:	3 "D" size batteries in series for reticle illumination and alignment function
Weight:	Body 380g (13.4 oz.) Optical fiber 150g (5.3 oz.) Camera adapter 280g (9.9 oz.)

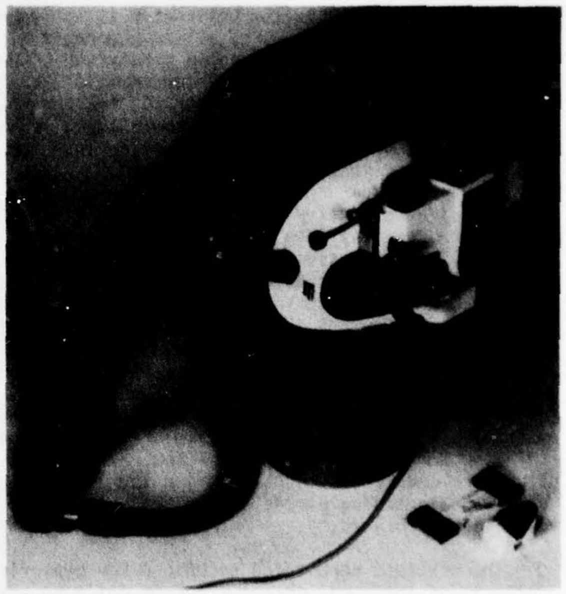


Figure 1. Basic NAC Eye Mark Recorder

The NAC basic structure is a hard plastic mask with pads which are adjustable by means of velcro tabs. The mask is secured to the face by nylon straps located over and around the head. An image lens which references the subject's field of view is located on the facial centerline and above the eyes. The scene referenced by this lens is transferred through the optical fiber bundle extending over the subject's head to a 16mm camera. The eye lamp is located slightly to the side and below the right eye. This unit provides the light source to the reticle which reflects off the eye. The half mirrors located in front of the eyes reflect the reticle through the Eye Mark optical recording path and superimpose the reticle on the recorded scene.

Adjustments and calibrations are accomplished by means of the xy adjusters, image lens knob, parallax adjustment knob, and Eye Mark focus dial. The xy adjusters are located on the mask next to the subject's left eye and provide xy fine adjustment for aligning the position of the Eye Mark. The image lens knob is located above the left eye and permits adjustment for the brightness of the field of view from T8 to T32. A parallax adjustment knob is located next to the right eye. This knob is provided to permit eye movement data acquisition in cases where short distance viewing (i.e., reading maps or books) is required. This adjustment is required only when the eye position exceeds a 15^{0} downward motion. An Eye Mark focus dial is provided and located above the right eye to permit correction for individual distance differences between the reflective mirror and eye position.

The recording adapter for the NAC is illustrated in Figure 1. The small end is directly interlocked with the optical fiber bundle while the opposite end utilizes a "C" mount ring to link the NAC system to a camera. The top of the adapter permits observation of the eye mark integrated into the visual scene and is also used to monitor adjustments and corrections. Adjustments and calibrations are accomplished by shining a beam of light through this recording adapter and moving the xy adjusters so that the 5mm diameter spot is superimposed on the cornea of the eye.

Camera. The camera arrangement consisted of a LOCAM Model 51-0002, high speed motion picture camera, with a Model 51-0088 BCD camera decoder; and a Model 13-0007 time code generator, all manufactured by Redlake Corporation. Specifications for the camera, BCD decoder and generator are presented in Tables 2, 3, and 4, respectively.

TABLE 2

General Specifications of the LOCAM Model 51-0002 Camera

Lens Mount:	"C" Mount with 0.690 + 0.0005 inches lens seat to film plane.
Electrical Connector:	The power connector at the rear of the camera is a 19 Pin Deutsch DM9601-19P connector.
Film Footage Indicator:	Displacement type footage counter, showing remaining film with scales for 100, 200 and 400 foot spools.
Operational Specifications	
Frame Rate:	Model 51-0002, infinitely variable between 16 and 500 FPS.
Frame Rate Stability:	± 1% or 1 frame, whichever is greater.
Frame Registration:	0.00013 Standard deviation.
Environmental Limits:	a) 60,000 ft. altitude b) -65 degrees F. (with heaters) c) Vibrations to 10g in all axes d) Operates under conditions of 95% RH.
Acceleration:	At maximum frame rate, the camera will not consume more than 10 + 2 feet of film in acceleration or deceleration.
Film Specification:	This camera accepts 16mm acetate-based films perforated for 0.3000 inch long pitch or 0.2994 inch short pitch. Thin-based films can also be used with optional film gate.
Film Capacity:	Camera accepts 100, 200 and 400 foot standard 16mm daylight loading spools.
Film Aperture:	0.292 X 0.410 inches with fiducials on one side and bottom.
Shutter:	Camera is furnished with a variable double disc shutter which can be manually adjusted from 0° to 160°. Shutters operate at 1/2 camera speed. For precise work below 20°, it is recommended that a single fixed blade shutter be utilized.
Power Source:	Model 51-0002: 28 +3, -0 volts DC
Maximum Operating Current:	Model 51-0002: 8.5 Amperes
Dimensions:	10.00 inches Lg. X 3.78 inches Wd. X 7.38 inches Ht. (excluding knob and lens; Ref: Fig 1-1)
Weight:	11 lbs. (excluding lens, film, and dovetail mount)

TABLE 3

General Specifications of the BCD Decoder Model 51-0088

Format:	4 x 9 matrix (9 digits BCD).
Input:	9 digits multiplexed BCD (parallel bit, serial digit).
Input Impedance:	50 Ohms (each line).
Print Sync:	Camera shutter correlation pulse commands display printout every frame during exposure.
Exposure:	*A four-position switch permits* selection of the adequate exposure for different film sensitivities.
Operation:	The recording head is mounted on the pressure plate, imprinting film through the base. Display is located on the right side (viewing film upside up) next to the image area, with the most significant digit on top and the most significant bit next to the frame.
Temp Range:	0C to 70C (32F to 158F), operational.
Package:	The on-camera decoder is housed in a 8 x 5 x 2 case mounted on the motor compartment cover plate of the LOCAM Camera.
Power:	The circuit is powered through the signal connector. 5V DC (4.75 to 5.25 V operational), .75 amp, nominal required.

TABLE 4

General Specifications of the Time Code Generator Model 13-0007

Output Format:	9 Digits (hrs, min, sec, millisec). BCD Code, multiplexed (parallel bit, serigal digit).
Panel Readout:	6 Digits (hrs, min, sec), Numeric LED display.
Accuracy:	1 PPM (.0001%)/-20C to 70C (-4F to 158F); 1 PPM (.0001%)/Month.
Sync Input:	Modulated or demodulated IRIG "B", at 0.5V to 20V pk. When present, this signal takes over timing control, which returns to internal time base when interrupted.
Time Code:	Selectable, actual IRIG time or arbitrary zero reference through time zero-reset input.
Other Codes:	Up to four digits (hrs and min places) can be set manually or by external BCD source (current syncing logic).
Temp Range:	0C to 70C (32F to 158F), operational.
Outputs:	2 parallel channels, up to 500 ft. of 50 to 500 ohm cable, unbalanced, coaxial or twisted pairs, at TTL levels.
Package:	6 x 6 x 9 in. drawn can with rubber lid seal.
Power:	28 VDC (21 to 30 VDC operational), 2.5A max. Provision for external battery input with automatic (internal) switching.

Figure 2 illustrates the NAC/camera arrangement. The LOCAM camera with BCD decoder is located to the far left of the picture. The recording adapter links the NAC recorder to the camera. Directly behind the camera is a (30 Vdc) battery supply which provides power for the time generator located to the right of the NAC. The smaller box is a modified power supply for the NAC and will be discussed later. The camera allows NAC data to be recorded on 400 foot, 16mm film at 16 frames per second. This footage is equivalent to approximately 15 minutes of data collection time. The time generator permits each frame of film to be coded with subject number, subject run, and real time in hours, minutes, seconds and milliseconds. Ease of handling and quick disconnections allow cameras to be switched with only two minutes of down time between subject runs.

Film and Lighting. As with any form of photography, the type film and lighting are critical items for favorable NAC data collection. In 1973, the laboratory realized the limitations of video tape (Figure 3) and switched to Kodak, high speed, daylight, Ektachrome E.F. color motion picture film (ASA 160/400 ft x 16mm). The film was exposed and processed by normal procedures. Although the use of film was a decided improvement over video tape, two shortcomings were still noted when this technique was used. First, contrast problems were observed when attempts were made to photograph both inside and outside the aircraft; and second, a critical illuminance problem limited data collection operations to about three hours per day. To correct these deficiencies, the laboratory converted to Kodak 4X negative black and white film (ASA 500/400 ft. X 16mm) in March 1975. The film was processed by developing a working positive print on Eastman reversal black and white print film (Figure 4). This change in film did not resolve the contrast issue but did extend the useful hours for photographing.

Finally, in June 1976, by utilizing Kodak high speed, daylight 434 7241 E.F. Ektachrome color film (ASA 160/400 ft x 16mm) and a light blue template around the instruments, the laboratory was able to film both inside and outside the aircraft as shown in Figure 5. To achieve these results, the film had to be exposed and processed at ASA 640.

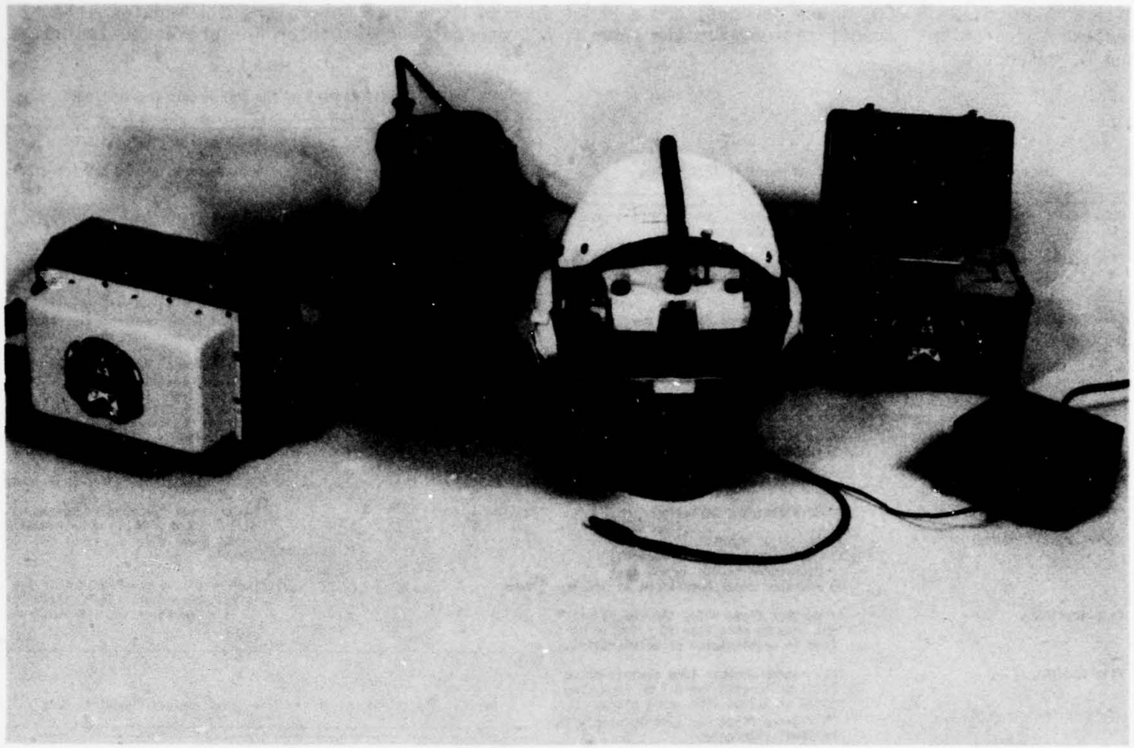


Figure 2. NAC/Camera System Arrangement

Figure 3. Video Tape of NAC Data

Figure 4. Black & White Motion Picture of NAC Data

Figure 5. Color Motion Picture of NAC Data

Through use of this film and special processing techniques, some guidelines for collecting favorable data were established. First, data collection could be accomplished whenever the sun angle was greater than 50° above the horizon. Filming could be accomplished in any direction except within 10° directly into the sun, at which time a "wash out" effect occurred. (*This rule was not restrictive to flights below the horizon within two hundred feet absolute*). Finally, after correcting for light degradation caused by deterioration of the optic bundle, a table was devised to determine the correct "T" stop for each luminance level. The bundle used by the laboratory had an approximate 3.3 T-stops light loss. Light level readings were derived from the average reflective light off the instrument panel and below the horizon. The values to determine each "T" stop for the NAC are referenced in Table 5.

TABLE 5

USAARL's "T" Stop Scale for the NAC Recorder

	Panel (Ft. L.)	Below Horizon (Ft. L.)
T8	11.75	31.4
T11	23.55	62.8
T16	47.10	125
T22	93.90	251
T32	188	753

HIMS. In addition to the NAC recorder and camera system, in-flight performance measures of psychomotor performance as well as aircraft performance were obtained via the USAARL Helicopter In-flight Monitoring System (HIMS). The HIMS measures pilot's cyclic, collective and pedal inputs and simultaneously records the aircraft responses to include position, acceleration, and rate changes. Twenty channels of continuous information are recorded in real time on an incremental tape recorder. The recorded values are then reduced and analyzed on ground based digital computers. A full report of the system is available in USAARL Report No. 72-11.[12]

The time code generator of the BCD camera decoder was synchronized with the internal time code generator of the HIMS. This time was printed on each frame of film as well as each second of performance data from the HIMS, thus visual data was synchronized with performance data.

Visual Free Time Chart. A visual free time chart similar to that utilized by Bell Helicopter Corporation[13] was designed to assist in determining pilot visual time which was not essential for aircraft management. The chart measured 14 X 9.5mm with 2mm lettering (Table 6). The words selected were random, meaningless, and monosyllabic. The chart was positioned on the instrument panel of the JUH-1 helicopter within visual range of the subject pilot.

Data Analysis Equipment. Equipment utilized for the reduction and analysis of the NAC data is listed in Table 7. The function of all equipment referenced is according to manufacturer's specifications.

TABLE 6

Visual Free Time Chart

feed sly as badge gape wrath pun cloth sick love rough kept calf
Greek beck nigh flop roe thick best fall choose flap jag frock chop
wasp true cheat tongue ode pass wink hitch hull browse zone kill
bog fee punt odds rooms lag shove kid fowl thigh hill trade bind
reap chart black scare writ wait high mast wife cob rind fling rot
pipe clothes mash vase good gage eyes rode lend forge raise sniff
puff yawn prime deep inch watch scan shank bronze thud grope
ray solve tug sup gap bathe curse slouch crib add owls thus clod
pus rear nose prig eat shine grudge flick dad gasp by wheeze
bored woo am roll slide though nine look ease act wire freak
queen dwarf aim spice jell scout shaft hum forth sledge south woe

TABLE 7

Data Analysis Equipment

	Company	Item	Model
1.	L&W	16mm Variable Speed Data Analyzer	Model 224A
2.	Hewlett Packard	Preset Counter	Model 5330B
3.	Hewlett Packard	Digital Voltmeter	Model 3480B
4.	Wang	Advanced Programing Calculator	Model 720
5.	Wang	Dual Magnetic Tape Cassette Reader/Recorder	Model 709
6.	Wang	High Speed Printer	Model 721
7.	Wang	Micro Interfacer	Model 705
8.	Wang	Input/Output Writer	Model 711
9.	Wang	Disc Drive	Model 710
10.	USAARL	12 Micro Switch Keyboard	

Modifications. The equipment shown in Figure 2 was compatible, as designed by the manufacturer, to document oculomotor performance in a laboratory setting. However, from the experience of the Aviation Psychology Division of the US Army Aeromedical Research Laboratory (USAARL), several modifications to the basic system were accomplished to provide a compatible system with the helicopter flight environment. Among the primary issues which had to be resolved were: calibration of the system; protection of the optic bundle; comfort of the mask for the subjects; and added stability for the mask to prevent misalignment of adjustment caused by the extreme vibration of the helicopter.

The modifications which were made to the mask to increase stability and provide more comfort for the subject are illustrated in Figure 6. The basic mask provided a padded "V" nose piece and nylon straps attached to three points on the mask which were at the top and on either side of the mask. This laboratory's modification was to remove the nose "V" and attach additional padding along sharp metal edges of the mask located along the forehead and bridge of the subject's nose. These locations were the most common points

of discomfort caused by the mask. By attaching additional snaps on the mask above each of the existing side snaps, the mask was adjusted to assume an improved, flatter position on the subject's forehead. This allowed the lower straps to fit under the occipital protuberance of the skull. The second set of straps was attached at about the Lambda and crossed laterally above the ears to the mask. The last strap started at the Lambda and ran anteriorly over the sagittal suture and frontal bone to the mask. These improvements relieved subject discomfort and provided needed stability.

The modification to the mask configuration was completed by the integration of a Gentex PBH2 tanker's helmet (Figure 7). Because the PBH2 helmet differed from the SPH4 aviator's helmet, in that it has a higher forehead protection plate, the PBH2 could be worn with the NAC recorder, while simultaneously affording head and ear protection in addition to increased stability for the total system. Additional nylon straps with snap connectors were attached to the top and back of the helmet to secure and protect the optic bundle.

Figure 6. Modifications to NAC System

Figure 7. Gentex PBH2 Tanker's Helmet

A variable power supply was designed and fabricated by the laboratory to aid in optimizing the maximum continued 5mm reticle from the eye lamp. The power supply utilized 28 Vdc aircraft power and provided a variable output of 1 to 5 Vdc. This unit supplanted the NAC recorder's normal "D" cell battery pack thus decreasing the risk of battery deterioration and power loss. In addition, an assortment of half mirrors with varying reflective characteristics was utilized to provide the best reticle possible for different luminance conditions.

To secure the LOCAM camera in the USAARL JUH-1 helicopter a special mount was fabricated from heavy gauge aluminum and fastened to the back of the subject pilot's seat. The camera was attached to the mount with a Winter dovetail camera mount and receiver. This mounting system provided a more compact arrangement of system components, aided in protecting the optic bundle by moving the total camera/NAC system with any movement of the seat, and insured system stability during flight.

The LOCAM camera was modified by the addition of an 8 amp fast blow fuse to the system to protect the diodes and camera motor. The 8 amp fuse provided the camera protection from damage caused by voltage transients or film jam. The fuse is located on the back of the main camera.

The final modification of the data collection system was the design and addition of a blue template which was placed over the existing instrument panel of the JUH-1 helicopter. A normal US Army instrument panel color was flat gray while this laboratory's was flat black. Neither panel allowed for proper contrast to photograph both inside and outside the aircraft. Through the use of a light blue template, this contrast problem was decreased to a minimal level.

PROCEDURE

Data Collection. The primary concern during the data collection phase was to assure the proper fitting of the NAC mask and the calibration/stabilization of the system. Without proper fitting the mask would cause discomfort within five to ten minutes of flight.

Initial fitting and calibration were performed in the laboratory with the subject seated and with numerous targets located in front of him to aid in calibration. After the basic mask and straps were fitted as previously discussed (refer to Figure 6), and the xy adjustment knobs were centered, a small pen light was used to "bore site" the crude adjustment of the NAC mask (Figure 8). The procedure was to focus the penlight so that its light would shine through the optic bundle receptacle of the mask, causing a 5mm dot to appear somewhere on the subject's right eye. The mask was then shifted so that the dot appeared directly in the center of the pupil of the right eye. After securing the mask for stabilization and comfort, the PBH2 helmet was fitted starting from the back and pulling it forward over the head. Once the helmet was on, chin strap fastened, and the mike boom adjusted, the mask was "bore sighted" again to assure that no significant changes in calibration had occurred.

The optic bundle was then connected and secured to the helmet and the recorder adapter added to the system. Again by placing the pen light in the monitoring section of the recorder adapter and by manipulating the xy adjustment knobs, the 5mm dot could be centered on the pupil of the right eye. Finally, after connecting the eye lamp to the NAC power supply, the NAC recorder was fine adjusted by the normal procedures outlined in the instruction manual. The advantage of this modification to the procedure was that the investigator could fit and calibrate the NAC in a relatively short period of time and the modification provided maximum handling protection to the optic bundle, a critical item of the system.

After calibrating the NAC, and before proceeding to the helicopter, the recording adapter was removed from the bundle to facilitate ingress of the subject pilot. Once seated, the normal safety procedures of fastening restraints and checking communications were accomplished. The adapter was then reconnected to the bundle, the total system connected to the camera, and a fine adjustment of the NAC performed.

Before starting a test profile, the helicopter was hovered for three to five minutes to allow the NAC time to "settle." This time was utilized to move the aircraft from the parking location to the test starting point. At this point, the NAC was fine adjusted a final time. After adjustment, a changeover ring on the recorder adapter was switched from the viewing position to recording and taped so that aircraft vibration would not cause slippage of the ring. After the test run was completed the reticle image was again checked to assure that no slippage had occurred during the flight. Previous studies have indicated that the corneal reflection technique has an overall accuracy of $\pm 2^0$ when used within $\pm 12.5^0$ of eye movement. Studies also indicated that head movement had to be considered any time eye movement was greater than 15^0 in any direction from the centrofoveal position.[1,8]

In an attempt to delineate the possible impact of errors caused by head rotation during flight, data were recorded on pilot head rotation during helicopter flight and the effects of this rotation on the visual data were assessed (USAARL Report No. 74-7).[8] The results of this study and subsequent data indicate that not only was head rotation during laboratory performance testing quite different from that observed in the operational environment; but also that head rotation presented minimal error in the collection of visual performance data during actual helicopter flight.

To facilitate scoring of the data, the aircraft windscreen in front of the subject was partitioned, using a grease pencil, into four equal sections. Twenty-five millimeter tick marks were added to the division lines, which allowed the data depicted on different locations of the windscreen to be recorded and distances outside the aircraft to be calculated. During the actual test profile, the changeover ring of the recorder adapter was utilized to segment portions of the film. By varying the changover ring from recording to viewing, three to four frames of film were overexposed, thus providing a reference mark delineating the start of each separate film segment. Finally, to complement the completed system, the visual free time chart was stationed on the instrument panel directly in front of the subject pilot. Although the NAC recorded total visual performance, the chart allowed the investigator to calculate the total free time in which the subject pilot was not utilizing critical cues to perform his assigned flight mission.

With a crew consisting of safety pilot, HIMS operator, photographer, and flight coordinator, each subject pilot was directed to perform any helicopter maneuver under study. The subject's total visual performance was recorded in real time on (motion picture) film. After the film was processed, these data were ready for reduction and analysis.

Data Reduction. The data reduction room was arranged so the investigator could utilize the L&W 16mm analyzer to project the NAC film on a screen. As shown in Figure 9, the film was processed at one quarter speed or four frames per second. Through a prereduction preparation, the total area of any one frame was divided into 13 major viewing areas. As the reticle appeared in each of the 13 areas the investigator pressed the corresponding microkey on a switch keyboard located on the table in front of him. The micro-switches were connected to the preset counter and the digital voltmeter to give each switch a selected voltage and an elapsed time for which the voltage selected was maintained. This information was transferred to the Wang disc drive through the microinterface and was stored.

Figure 8. Crude Calibration of NAC System

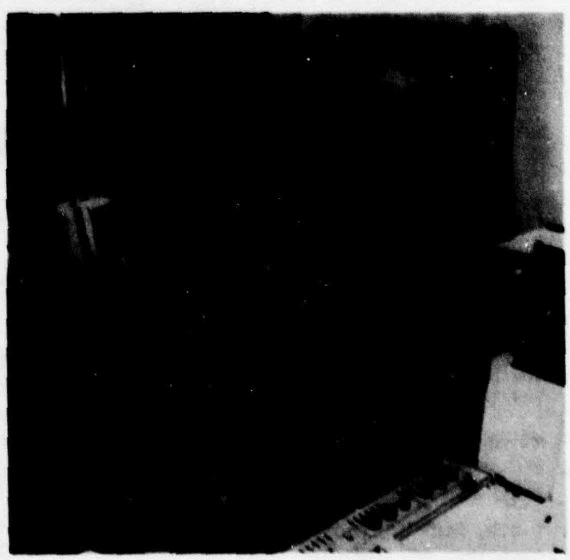


Figure 9. Data Reduction Using Projector and Wang Computer

After position and time data from the film had been stored on the disc, the data were then transferred to cassette tapes so that each individual flight segment of data could be selected, values computed, and results printed by the high speed printer. The final results gave the total time elapsed in each of the 13 zones, percent of the total time, frequency, mean time in each zone, and the standard deviation along with link values from each zone to all others. After each segment was computed, a master total was summed for the total film. The time for data reduction was approximately five hours for each one hour of data.

APPLICATION

To date, visual performance/workload has been recorded via the USAARL method during simulated instrument flight (IFR), helicopter instrument flight, helicopter visual flight (VFR), copilot/navigator nap-of-the-earth and pilot terrain flights. In-flight visual performance data on 53 aviators of varying experience levels (i.e., 200 to 2500 flight hours) have been obtained comprising approximately 30 hours of real time film recordings.

The simulated instrument flight and the helicopter instrument flight data have been reduced and are presently undergoing analysis. For the simulator study, preliminary results utilizing multivariate analysis techniques indicate that the experience level of the two test groups (i.e., 200-hours flight experience and 1500-hours flight experience) was not a contributing factor to the percent of viewing time, or frequency of use for particular instruments. For both the simulator and in-flight investigations, subjects utilized the vertical situation indicator and the horizontal situation indicator approximately 43 percent and 26 percent of the time, respectively. The next most frequently utilized instrument comprised only 8 percent of the total viewing time and the total arrangement of the seven monitoring gauges (i.e., oil pressure and temperature) required only 2 percent of this time. In addition, it was found that the mean time spent on more complex indicators such as the vertical situation indicator (.5 to .6 sec) was considerably longer than that spent on one and two pointer instruments (.2 to .3 sec).

Results of the five studies utilizing the NAC Eye Mark recorder to study visual performance/workload of Army pilots during helicopter operations will be published in future USAARL technical reports.

DISCUSSION

One of the usual techniques for determining visual performance/workload is to solicit opinions from aviators, asking them to describe those areas which they feel provide the necessary visual cues for safe helicopter flight. However, there is an extremely low correlation between aviator opinion and objective visual performance data when these two techniques are utilized simultaneously.[8,14] Indeed, it would seem that experienced aviators may be responding to cues which have become useful through experience and not necessarily those which they were taught and which they claim are those which provide them their best visual information. Therefore, continued objective research data are required to isolate those visual cues critical for helicopter flight.

The methodology outlined in this work provides an effective means of measuring visual performance/workload through the corneal reflection technique. Utilizing the NAC Eye Mark Recorder, camera system, and the helicopter in-flight monitoring system (HIMS), data have been collected relating visual psychomotor and aircraft performance during multimission operations. The significance of this research resides in its potential to provide useful information about aviator sensory and motor performance in the rotary wing environment which may be applied to problems in current instrument panel designs and provide aids for innovative and efficient instrument displays for future rotary wing aircraft.

DISCLAIMER

The findings in this report are not to be considered as an official Department of the Army position unless so designated by other authorized documents.

REFERENCES

1. Hall, Robert J. and Cusack, Bruce L. The measurement of eye behavior: critical and selected reviews of voluntary eye movement and blinking. Technical Memorandum 18-72, July 1972, Human Engineering Laboratory, Aberdeen Proving Ground, Maryland.

2. Klein, Richard H. and Jex, Henry R. An eye-point-of-regard system used in scanning and display research. Paper presented at the SPIE 15th Annual Technical Symposium, Anaheim, California, September 1970.

3. Monty, Richard A. An advanced eye-movement measuring and recording system, American Psychologist, 1975, Vol 30 (3), 331-335.

4. Mowrer, O. H., Ruch, R. C., and Miller, N. E. The cornea-retinal potential difference as the basis of the galvanometric method of recording eye movements. American Journal of Psychology, 1936, Vol 114, 423.

5. Milton, J. L, Jones, R. E., and Fitts, P. M. Eye fixations of aircraft pilots: frequency, duration, and sequence of fixations when flying the USAF instrument low approval systems (ILAS). USAF Technical Report No. 5839, October 1949. Wright Patterson Air Force Base, Dayton, Ohio.

6. Fitts, Paul M., Jones, Richard E., and Milton, John L. Eye fixations of aircraft pilots: frequency, duration, and sequence fixations when flying Air Force ground controlled approach system (GCA). USAF Technical Report No. 5967, November 1949. Wright Patterson Air Force Base, Dayton, Ohio.

7. Mackworth, N. H. and Thomas, E. L. Head-mounted eye marker camera. Optical Society of America Journal, 1962, Vol 52, 713-716.

8. Frezell, Thomas L., Hofmann, Mark A., and Oliver, Richard E. Aviator visual performance in the UH-1H, Study I. USAARL Report No. 74-7, October 1973. U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL.

9. Frezell, Thomas L., Hofmann, Mark A., Snow, A. C., Jr., and McNutt, Richard P. Aviator visual performance in the UH-1H, Study II. USAARL Report No. 75-11, March 1975. U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL.

10. Young, Lawrence R. Measuring eye movements. American Journal of Medical Electronics, 1963, 300-307.

11. Sanders, M. G. Personal communication based on data being prepared for publication. Nov 1976.

12. Huffman, H. W., Hofmann, M. A., and Sleeter, M. R. Helicopter in-flight monitoring system. USAARL Report No. 72-11, March 1972, U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL.

13. Strothers, D. D. and Hartfield, H. W. Improved cobra armament program, Human Factors Engineering Final Report No 209-099-381, August 1973. Bell Helicopter Company, Fort Worth, Texas.

14. Simmons, Ronald R., Hofmann, Mark A., and Lees, Michael A. Pilot opinion of flight displays and monitoring gauges in the UH-1 helicopter. USAARL Report No. 76-18, April 1976. U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL.

DISCUSSION

LOVESAY: (United Kingdom) I would like to congratulate the speaker on the excellence of his Eyemark film. Did the pilot wearing the Eyemark camera know the area flown over and did he have to navigate by map?

SIMMONS: (United States) Pilots were given a variety of tasks to perform and areas over which to fly. There was no need to map-read. The on-board safety pilot acted as the navigator as needed and instructed the pilot on where to go as the flight progressed.

SANDERS: (Netherlands) How did you control for head rotations during flight? What effects did they have?

SIMMONS: (United States) The camera did not move. Only the lens mounted on the mask moved. The image was transferred via an optic bundle to the camera. I should add that head movement in the real world environment is more frequent than that observed in the laboratory environment. Therefore, eye movement to extreme angles did not occur and thus no error due to this factor was introduced into the data. Finally, I should add that eye position signals were verified during the calibration procedures in the laboratory.

ADAMS: (United States) I noted in the film that the pilot made only 2 eye movements to the panel. The remainder of the time, he made fixations in a very narrow band directly ahead at the trees. Was this behavior typical? What did the pilot look at on the panel?

SIMMONS: (United States) Fixation outside the cockpit is typical. The infrequent fixations inside the cockpit are on the airspeed indicator. I feel it is unfortunate that the pilot during nap-of-the-earth (NOE) flight is not monitoring aircraft performance, nor is the copilot. Apparently, at these velocities and altitudes, pilots do need some velocity feedback and in clear areas they do spend more time fixating on the panel displays.

HOLLOWAY: (United States) Did any incidents or emergencies occur in the course of a flight which permitted you to observe eye movements under these stressful conditions?

SIMMONS: (United States) No incidents occurred secondary to the use of the apparatus. No emergencies occurred, but we are looking for eye movement data if such emergencies occur.

PERDRIEL: (France) Has your laboratory observed differences in visual performance during simulated helicopter flight compared to operational or "real world" flight? How does your laboratory provide for maximum safety during the research flights?

SIMMONS: (United States) We have compared data from the simulator and from operational UH-1 helicopter flights. Data from simulated flights are different from real world flights. The two sets of data don't correlate. You must go to real world flights to get data for operational applications. There is a continuous need for operational data. With regard to safety, we always provide a fully qualified safety pilot. It is made clear in preflight briefings that the safety pilot takes command in an emergency.

USE OF EYE-MOVEMENT MEASURES TO ESTABLISH DESIGN PARAMETERS FOR HELICOPTER INSTRUMENT PANELS

John A. Barnes
Engineering Research Psychologist
U. S. Army Human Engineering Laboratory
Aberdeen Proving Ground, Maryland 21005

SUMMARY

The helicopter in transition from a fair-weather aircraft to an all-weather machine has increased the pilot's workload substantially. For sometime the Systems Performance and Concepts Directorate's Aviation Team has been working towards eliminating, by design refinement, as much of this additional pilot workload as possible.

We felt that there must be a way to actually prove the "worth" of a particular instrument panel design. By recording the pilot's eye-scan paths and fixation points during actual helicopter flight, we have a method which provides us with an accurate measure of the visual workload imposed by a particular panel design. This tool is not limited to aircraft instrument panels; it can be used to evaluate any operator control panel or even the design of a multi-information display within a panel.

Our initial work was done in the UH-1 helicopter using experienced instrument-rated pilots flying actual maneuvers on instruments. The knowledge gained from these data allowed us to design a helicopter instrument panel in which the most referred to instruments were placed so that the eye-scan paths were minimized. This design considerably lessens the pilot's visual workload, cuts fatigue and allows him more time for other tasks.

We are also using actual flight eye-movement measures to evaluate the target detection tasks of the observer during nap-of-the-earth flight in the OH-58 scout helicopter.

INTRODUCTION

When attempting to evaluate an operator's performance in a task as complicated as a rotary wing pilot's we generally find that we can only measure the interface between the man and the task at hand, that is, his reactions to the information available. We know very little concerning what part of the available information was used to elicit the operator's response, thus we can only infer as to the existing stimulus-response relations. When information is presented both by an extremely complicated display such as the helicopter instrument panel and the part of the real world available through the wind screen, we have a task that appears to be confounded beyond solution.

The measurement of the pilot's eye movements and points of fixation while he is performing the flight task enables one to at least pinpoint one of the major sources of information used by him to perform this complex task. These measurement techniques allow one to determine the pilot's fixation point to within a few seconds of arc, the length of these fixations, and the paths from one fixation point to the next.

One of the earliest documented uses of eye movement technique to measure a pilot's workload was done in 1943[6] to determine the amount of pilot fatigue induced by two different instrument panel configurations. Since then there have been many studies of fixed wing pilot's eye movements using an equally large assortment of measurement devices.

The use of the pilots' eye movement data for the evaluation of instrument panels is not as straight forward as one might imagine, other experimenters[5] and the Human Engineering Laboratory (HEL)[1,3] have found that the point of greatest fixation time is not always the point from which the pilot is deriving the major portion of his information, rather the instrument that has the greatest frequency of fixations is the one that pilot considered most necessary to perform the task at hand. The point of greatest fixation time was often found to be a point from which it was possible to peripherally monitor instruments the pilot considered important. This point can also, if it is a particular instrument, be an indication of difficulty of interpretation of the instrument reading. The pattern of the eye movements, the paths, link values, between the instruments, are the direct indication of the visual work load imposed by that particular instrument panel design.

METHOD

HEL's first program of pilot's eye movement study was begun in 1969[1] to investigate the means by which a UH-1B helicopter pilot gained the information needed by him to control his vehicle in space. We used a Westgate EMC-2 Eye-Movement Camera, Figure 1, which is a helmet mounted camera system. We modified the system slightly to provide a more comfortable fit for the subjects and to stabilize the camera so that large head movements would not cause a loss of calibration during flight. This camera system uses a spot of light to mark the fixation point, Figure 2, which enabled us to determine the exact location of the fixation points. Other system designs which rely on manual eye angle calculations to assign the fixation point have a tendency to misassign fixation points which fall between instruments. The two hours of flight film data gathered for this program provided considerable information about what instruments the pilots used, how long they fixated on a particular instrument, their scan paths, links, between the instruments and the frequencies at which they referred to particular instruments.

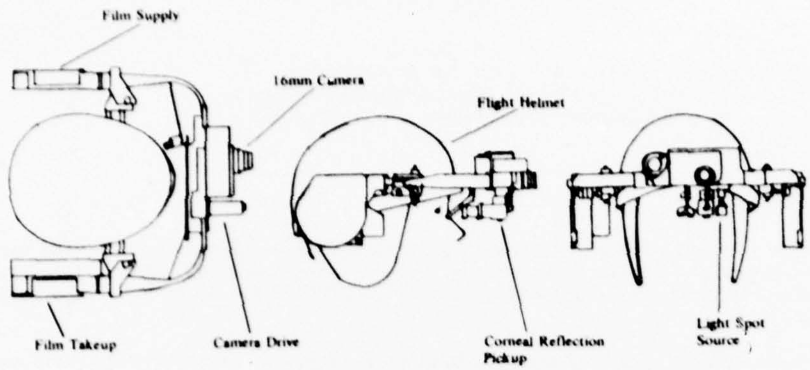


FIGURE 1. EYE-MOVEMENT CAMERA SYSTEM

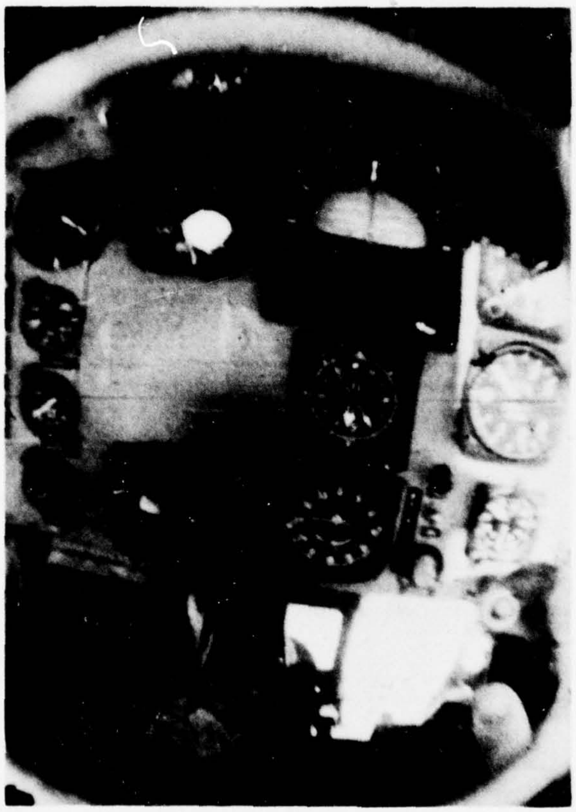


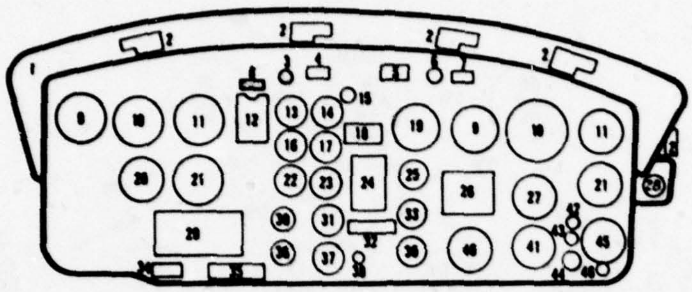
FIGURE 2. INSTRUMENT PANEL WITH FIXATION SPOT
(500 feet per minute climb)

RESULTS

Prior to the actual flight test a group of instrument qualified U. S. Army UH-1B helicopter pilots were asked to describe in detail what instruments they used to perform each of the 21 maneuvers that were flown in the flight test. The correlation of the interviews/questionnaires and the eye movement fixations and scan paths from the flight film helped strengthen our interpretation that peripheral vision was being used when the fixation points were between instruments.

The use of the UH-1B helicopter was a fortunate choice of helicopter for the HEL eye movement work because its instrument panel, Figure 3, was a good example of an instrument panel designed to comply with the existing military specifications and standards and it was an instrument panel that was quite acceptable to most helicopter pilots. The spacing of the instruments was such that we were easily able to determine

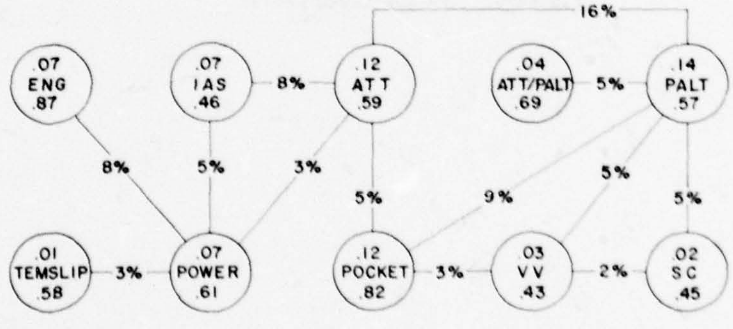
that the pilots were fixating on blank portions of the panel rather than these seemingly useless fixations being caused by loss of calibration in the eye movement measurement system. The panel had several non-instrument fixation points that were used frequently; they were: a point located midway between the Attitude Indicator, Remote Magnetic Indicator, Altimeter, and a Vertical Velocity Indicator which we called Pocket; a point located between the Airspeed Indicator, Dual Tachometer, and the Torquemeter which we called Power; a point located between the Gas Producer Tachometer, Exhaust Gas Temperature Indicator, and the Turn and Slip Indicator which we called Tempslip. The pilots would also fixate on points between the Attitude Indicator and Airspeed Indicator, between the Altimeter and the Vertical Velocity Indicator, and between the Attitude Indicator and the Altimeter.



1. Glare Shield
2. Secondary Lights
3. Engine Air Filter Light
4. Radio Call Designator
5. Master Caution Light
6. RPM Warning Light
7. Fire Detector Test Switch
8. Fire Warning Indicator Light
9. Airspeed Indicator
10. Attitude Indicator
11. Altimeter
12. Compass Correction Card Holder
13. Fuel Pressure Indicator
14. Fuel Quantity Indicator
15. Fuel Gage Test Switch
16. Engine Oil Pressure Indicator
17. Engine Oil Temperature Indicator
18. Cargo Caution Decal
19. Dual Tachometer
20. Radio Magnetic Indicator
21. Vertical Velocity Indicator
22. Transmission Oil Pressure Indicator
23. Transmission Oil Temperature Indicator
24. Pilots Check List
25. Torquemeter Indicator
26. Go-No-Go Take-off Data Placard
27. Radio-Magnetic Indicator
28. Standby Compass
29. Operating Limits Decal
30. Main Generator Loadmeter
31. DC Voltmeter
32. Engine Caution Decal
33. Gas Producer Tachometer Indicator
34. Engine Installation Decal
35. Transmitter Selector Decal
36. Standby Generator Loadmeter
37. AC Voltmeter
38. Compass Slaving Switch
39. Exhaust Gas Temperature Indicator
40. Turn and Slip Indicator
41. Omni Indicator
42. Marker Beacon Light
43. Marker Beacon Volume Control
44. Marker Beacon Volume Control
45. Clock
46. Cargo Release Armed Light

FIGURE 3. UH-1B INSTRUMENT PANEL

Figures 4 through 7 are from an HEL Report[3] which compared helicopter and fixed wing eye movement studies. Figure 4 is the link diagram of UH-1B straight and level flight, each circle represents an instrument/fixation point, the number at the top of the circle is a portion of the total time that the instrument/fixation point was fixated during the maneuver, the number at the bottom of the circle is the average fixation time in seconds.



LINK VALUES BASED ON 4 FLIGHTS VALUES LESS THAN 2% OMITTED

FIGURE 4. ROTARY WING STRAIGHT AND LEVEL

The abbreviations used for instrument identification in these figures are:

ATT	Attitude Indicator
IAS	Airspeed Indicator
PALT	Altimeter
ENG	Engine Instrument Group
RMI	Remote Magnetic Indicator
VV	Vertical Velocity Indicator
SC	Standby Magnetic Compass

The lines that connect the various instrument/fixation points are the links or scan paths and the numbers indicate the percentage of the total scanning activity that took place between these two instrument/fixation points. Figure 5 is the link diagram for the UH-1B standard rate turn;

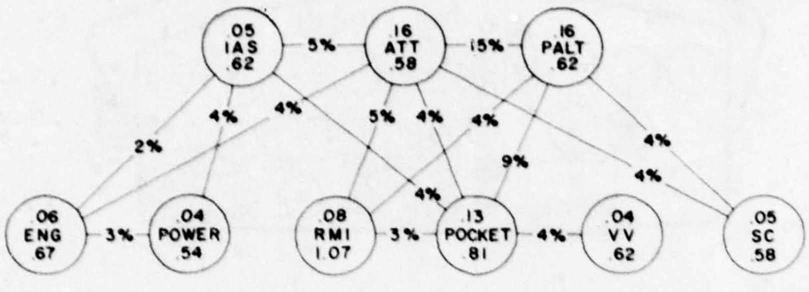


FIGURE 5. ROTARY WING STANDARD RATE TURN

this maneuver caused the pilots to add the RMI as a fixation point but notice the very slight change in the use of the peripheral vision fixation point called Pocket. Figures 6 and 7 are link diagrams of maneuvers that were not a planned part of the test scenario, but one of the pilots was a skilled instrument pilot and flew these two maneuvers totally on instruments. The increased difficulty of the maneuvers is apparent in the increase in the number of fixation points and in the number of links and the decrease in their values.

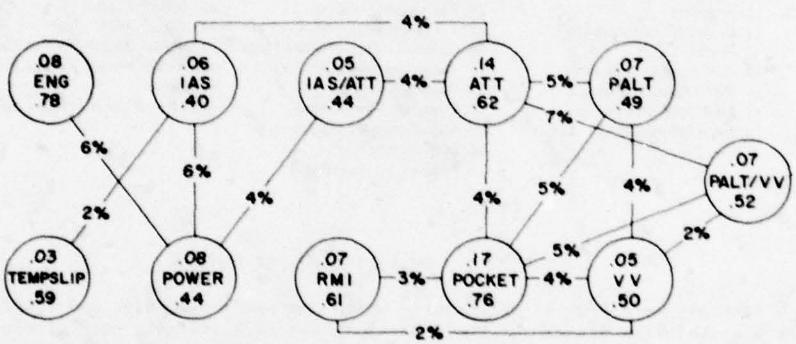


FIGURE 6. IFR LOW LEVEL CRUISE

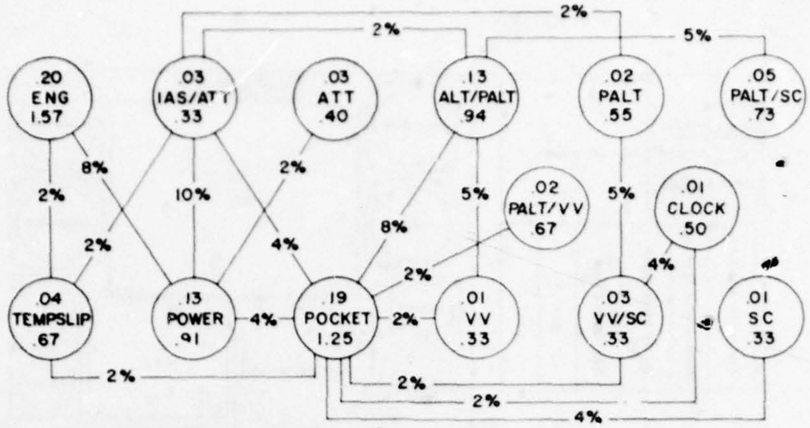


FIGURE 7. IFR HOVER IGE

The point of greatest fixation time and repeated points of fixation which are points between two or more instruments are the major clues for instrument panel designers. As stated previously, these fixations indicated that the pilot was using peripheral vision to monitor adjacent instruments which were important to the maneuver being performed. This use of peripheral vision is well known, yet, to our knowledge, only one other document[5] on eye movement measurement has reported on this use of peripheral vision by pilots to monitor several instruments from a single fixation point. This point or points may move as the maneuvers change, or if the instrument panel design is very poor there will be no peripheral fixation points. When the panel design is poor the frequency of change of fixations, link values, between the various instruments will increase. This increase in link values is also an increase in pilot visual work load. The goal of the instrument panel designer is to keep the pilot visual work load, link values, as low as possible with the ultimate goal being a single display to replace the total panel.

APPLICATIONS

The information gained from the eye movement flight data was applied to the design of a new helicopter instrument panel.[2] This instrument panel was to be a part of a crew station design for a twin engined U.S. Air Force, Army, Coast Guard, and Navy Rescue/Utility helicopter. The 54.25 x 14 inch (133 x 36 cm) was designed to use conventional "off-the-shelf" aircraft instruments where possible. Figure 8 is a picture of the mockup of that panel, Figures 9 and 10 are annotated drawings of the panel. The only item that was not available at the time was the multimode cathode ray tube (CRT) Vertical Situation Display (VSD) for the pilot. This display, as depicted, was within the state of the art and encompassed the needs set forth by the four services in one unit. It was not manufactured in this form as no one service required all of the functions depicted. It was to have had an analog "Shades of Gray" Terrain Following (TF) mode, an analog Instrument Landing System (ILS) mode, an analog Flight Director (FD) mode, an analog Hover Indicator (HI) mode and an infrared (IR) television mode. The multimode VSD was intended to be the central fixation point for the pilot with the remaining flight instruments placed immediately adjacent to it.

FIGURE 8. RESCUE/UTILITY HELICOPTER INSTRUMENT PANEL

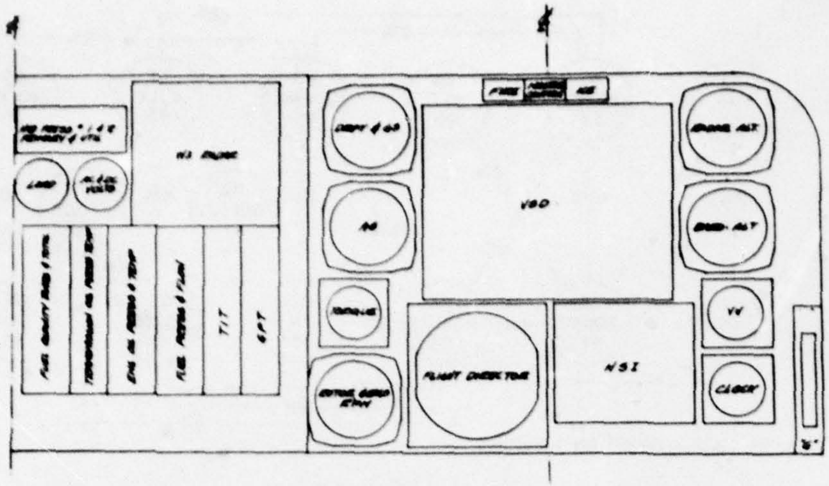


FIGURE 9. PILOT'S INSTRUMENT PANEL

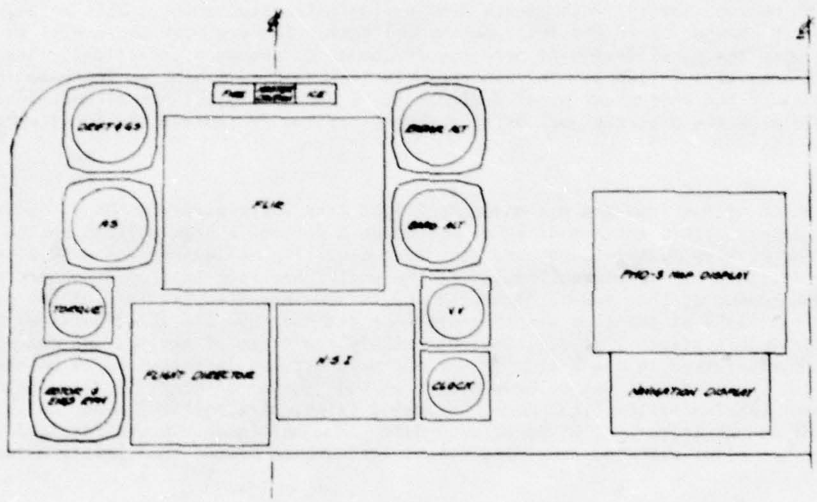


FIGURE 10. NAVIGATOR/COPILOT'S INSTRUMENT PANEL

A conventional Flight Director was placed directly below the VSD so that in case of a failure, or when some mode other than FD was selected, the Flight Director attitude information would be available to the pilot with a minimal change in the length of his scan paths to the other flight instruments. If there was to be a fixation point on this panel like the "Pocket" on the UH-1B, the design would indicate that it would be at the base of the VSD between the Flight Director and the Horizontal Situation Indicator; from this point normal peripheral vision would encompass all of the flight instrument group except the Radar Altimeter and the Drift and Groundspeed Indicator. The Radar Altimeter was placed at the top of the panel because, for the normal mission of this helicopter, it would be the least used of the three altitude related instruments. The same rationale was used for the placement of the Drift and Groundspeed Indicator. The instrument panel, Figure 11, used in one of the current competition utility helicopters for the U. S. Army generally follows this design. A drawing of the advanced version of the Rescue/Utility helicopter instrument panel is shown in Figure 12. This instrument panel features vertical scale instruments which, while not "off-the-shelf" items, were available on special order from one or more instrument manufacturers. Some of these instruments were moving tape fixed indicator types and others were moving indicator fixed tape design. This panel could impose a greater visual work load on the pilot in that the pointer position clue is not available on the moving tape instruments; they have to be "read" to achieve information transfer. Fixed wing simulator studies using moving tape instruments[5] do not indicate any increase in visual work load but as yet we do not have any simulator or actual flight eye movement data on helicopter flight using these instruments.

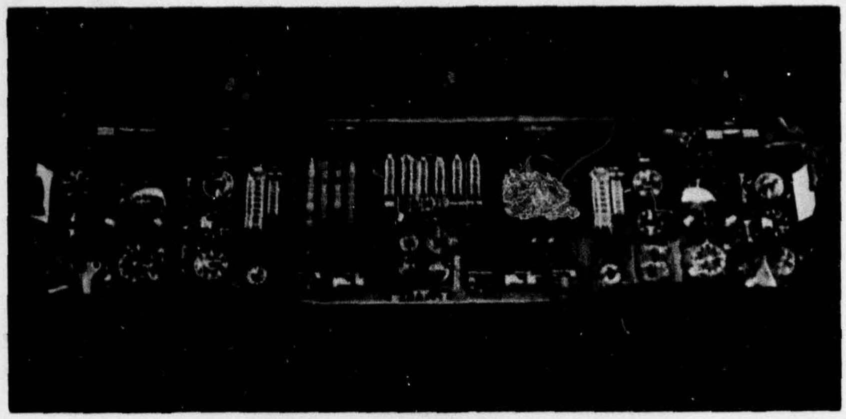


FIGURE 11. UTILITY HELICOPTER INSTRUMENT PANEL (PILOT).

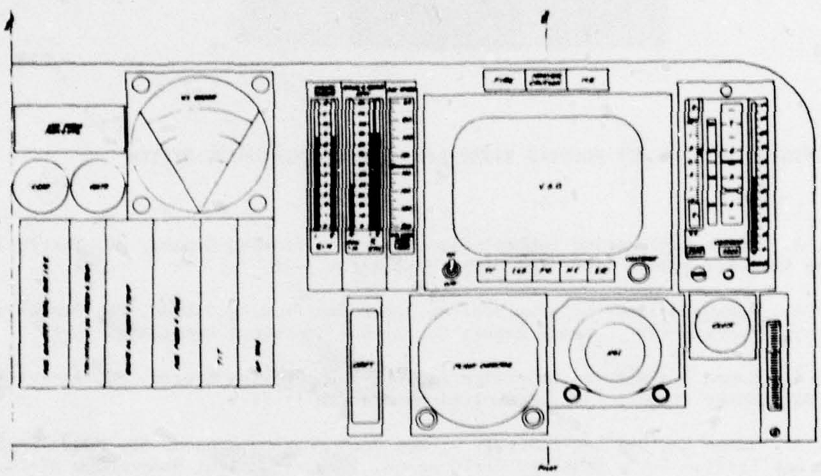


FIGURE 12. ADVANCED RESCUE/UTILITY HELICOPTER INSTRUMENT PANEL

The data from the initial flight tests have also been used to produce a computerized design program for helicopter instrument panels.[7]

CURRENT AND FUTURE CONSIDERATIONS

Our current eye movement work[4] has been concerned with the air-to-ground target detection performance of helicopter observers flying at nap-of-the-earth altitudes. The new eye movement measuring system, Figure 13, which we developed from "off-the-shelf" components provides us with a video record of the subject's eye movements with the fixation point indicated by a spot of light superimposed on the real world scene the subject is viewing. This system is mounted on a standard aviator's helmet and weighs 15 ounces (430g). The subject is allowed complete and unrestrained head movement and his body movement is limited only by the length of the wire from the video output unit to the monitor/recorder. We intend to use this system to evaluate new multi-element displays, control panels, and instrument panels of new air and ground vehicles. The system will also be used to evaluate operator performance in air and ground tasks.

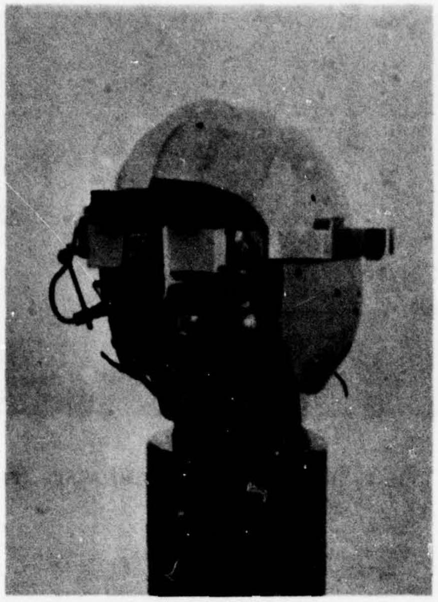


FIGURE 13. HELMET MOUNTED VIDEO EYE MOVEMENT MEASURING SYSTEM

REFERENCES

1. Barnes, J. A., Human Engineering Laboratories, Aberdeen Proving Ground, MD, Tactical Utility Helicopter Information Transfer Study, 1970, Technical Memorandum 7-70.

2. Barnes, J. A., Human Engineering Laboratories, Aberdeen Proving Ground, MD, Search and Rescue Crew Compartment Mock-Up Program, 1971, JANAIR Report No. 710601/Technical Memorandum 16-71.

3. Barnes, J. A., Human Engineering Laboratories, Aberdeen Proving Ground, MD, Analysis of Pilot's Eye Movements During Helicopter Flight, 1972, Technical Memorandum 11-72.

4. Barnes, J. A., Human Engineering Laboratory, Aberdeen Proving Ground, MD, Human Engineering Laboratory Camouflage Application Test, Observer Performance, 1976, Technical Memorandum 32-76.

5. Gainer, C. A. & Obermayer, R. W., Pilot Eye Fixations While Flying Selected Maneuvers Using Two Instrument Panels, Human Factors, 6 (5), 1964, 485-501.

6. McGehee, W., Comparative Study of Pilot Fatigue Resulting from Extended Instrument Flight Using Standard AAF and British Instrument Panels, U. S. Navy ATL-R601, 1943.

7. Wyman, S. D., Evaluation of Computerized Layout Algorithms for Use in Design of Control Panel Layouts, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia, 1975.

DISCUSSION

JAHNS: (United States)
Does the head-mounted equipment in any way alter the visual scanning behavior and/or the decision criterion for the target acquisition task?

BARNES: (United States)
Each of our subjects flew missions with and without the Eyemark system and the results were compared. There were no statistically significant differences in the ranges at which the targets were acquired or in the time it took to acquire them. We concluded that the device did not hinder the subject.

SANDERS: (Netherlands)
Were there any differences between the movement of the helmet and the movement of the head?

BARNES: (United States)
We think not. We use a molded helmet liner and a football chin strap and feel that the head and helmet are well coupled.

SANDERS: (Netherlands)
Are you concerned about what might be a discrepancy between your results and from a different field? You find a low correlation between subjective reports about what is seen and objective measures, but high correlations are reported in, say, advertisement research, between subjective and objective data.

BARNES: (United States)
I think there is a basic psychological difference in these two situations. The aviator is under a considerable workload when he is engaged in instrument flight and he does not assess accurately the amount of time he spends on each instrument. Therefore, he reports various observation times according to what he was taught would be appropriate. A person looking at an advertisement is not under a workload and has no previous instruction, and so he can report accurately on the content of the advertisement.

DOETSCH: (Germany)
In the previous paper, special effort is made to ensure that the eye movement camera follows faithfully the head movements. You have interposed the additional mass of the helmet in the kinematic chain. Does this not produce more errors between actual head movements and camera-recorded eye movements?

BARNES: (United States)
We use a special foam helmet liner molded to fit the individual subject and mounted to the Eyemark helmet by two bolts. We also use the football chin strap. As I said before, I feel this provides the stability necessary for the system to track faithfully.

SIMMONS: (United States)
My own experience is that you still have skin and hair movement. Have you been able to correct this?

BARNES: (United States)
The special foam liner we use has been equally as stable as the padded face mask and augmented head strap system used elsewhere. We have used both and have found that our system does an excellent job. I feel we have eliminated skin movement because of the way our system is anchored.

NICHOLSON: (United Kingdom)
Is it possible that there could be a neurophysiological basis for inadequate recall of the work pattern? Several years ago, I was involved in a study in which the pilot was required to assess his workload during the approach and landing of a transport aircraft. From this we developed a model for assessment of workload for this pilot. In a few landings, it was found that the model did not hold and these were high workload landings accompanied by high heart rates during the landing and gross finger tremor after landing. It was suggested that the high workload led to an arousal of the central nervous system above that which would be optimum for the task, and this was associated with impaired ability to assess the task. Could a similar explanation be useful in your studies?

BARNES: (United States)
None of the maneuvers our helicopters were performing were of a comparable level of workload as in the landings you refer to. Therefore, I doubt seriously if workload was involved. The answers given by the helicopter pilots when they were asked what instruments they used to fly a particular maneuver correlated highly with the "Book" answers for these maneuvers; their actual instrument use reflected instrument use patterns they had developed during many hours of actual instrument flight. Perhaps less experienced pilots would have followed the "Book" more closely.

AUDITORY COMMUNICATION AND WORKLOAD

by
Roger Green and Ray Flux
Royal Air Force Institute of Aviation Medicine, Farnborough, Hants

SUMMARY

This paper discusses the general problem of assessing how psychomotor workload may interfere with the performance of an auditory communications task. Two experiments are described. The first illustrates that an auditory task from which a cumulative response time measure is taken is affected by changes in signal quality and the second experiment shows that the same auditory task is also affected by a realistic form of workload (flying an aircraft simulator). The implications of these findings are discussed both in terms of assessing the quality required of a communications system and in terms of the use of secondary tasks in the assessment of workload.

INTRODUCTION

The measurement of the intelligibility of communications systems is a commonly undertaken procedure. In the military situation, the technique has been used in assessing all aspects and components of communications systems from microphones to headsets and helmets. The comparative assessment of such items is a task for which intelligibility tests are ideally suited, but the absolute level of intelligibility required in order for a system to be acceptable to the operator and to be operationally adequate is less easily defined.

Before any estimate can be made of the level of intelligibility required of a system, it is desirable to know what effects workload and stress are likely to have on the reception and understanding of speech. However, the overriding problem of experimental investigation is one of selecting a communications task or intelligibility test that does not require a written answer. All normal intelligibility tests require the listener to either write down a word which he hears, or to delete one of several possible words on a printed form. Neither of these responses is possible when the operator is required to perform concurrently some psychomotor task. The obvious alternative is to use spoken responses, but these are almost equally unacceptable because of the difficulty of marking - the intelligibility test has become two intelligibility tests.

The present experiment attempts to circumvent these response difficulties by utilising the fact that intelligibility can be assessed by measurement of response or recognition times. Pollack (1963) has shown that there is a relationship between intelligibility and response time to random, monosyllabic words, but his technique is unsuitable for use in the present experiment. Holloway (1972) however, developed a technique for measuring cumulative response times to spoken digits, and it was a method similar to this which was employed in the current experiment.

The main purpose of the first experiment, therefore, was to determine whether the type of approach developed by Holloway is one which can be used in assessing military communications equipment, and whether any unforeseen problems are encountered when using the technique at the same time as a concurrent psychomotor task - in this case, the step-tracking task devised by Gibbs.

METHOD

Apparatus

The speech communications system consisted of an intercommunications system into which defined levels of white noise could be injected. This was achieved by using two Lustrephone VC 80/HS headsets fitted with Lustrephone DI 59 microphones, connected through a Quad 303 amplifier. White noise was injected into only the subject's (listener's) headset via a Vortexion type H/3M mixer unit and the Quad amplifier. The signals from each microphone were also recorded on separate channels of a Ferrograph Series 7 tape recorder. The noise level at the ear of the subject was determined by measurement with a Bruel and Kjaer artificial ear and Type 2603 microphone amplifier.

The tracking task display was generated on the screen of a Digital Equipment Corporation 338 Visual Display Unit and relayed to the subject by closed circuit television equipment. The subject was also provided with a joystick which was connected to the computer which was driving the display permitting the subject to have control over the pursuit part of the tracking task.

The experimenter was also provided with controls enabling him to start the tracking task and equipment with which to time the audio task. The disposition of this equipment is illustrated in Fig 1.

The Tasks

The listening task was similar to that used by Holloway (1972). The subject and experimenter (who acted as a "pacer") were connected via an intercommunications system. The pacer read the first digit from a list of 50 digits to the subject who was required to add 3, as quickly as possible, on to the digit he heard and report the answer to the pacer. The pacer read the next digit from his list as soon as he heard any response from the subject, regardless of its accuracy. The only reason for the subject being required to add 3 on to the digit he heard was to slow down the pace of the experiment to a manageable level.

The white noise, heard only by the subject, was at one of four sound pressure levels, 64, 72, 77 or 85 dBA. These levels were arrived at subjectively, the lowest representing a scarcely audible level in the presence of the digits, and the highest being that level in which total recognition accuracy for the digits could be relied upon.

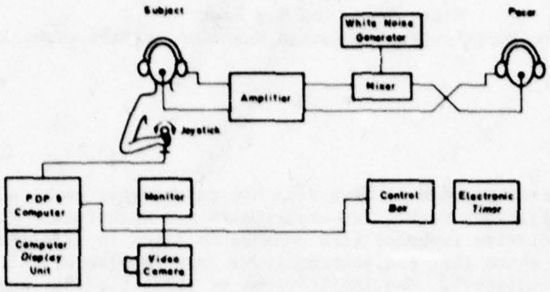


FIG. 1 Experimental Layout

The pacer could not hear the noise and was thus not induced to modify his speech level, which he attempted to keep constant by monitoring the VU meter on the tape recorder. This speech level corresponded, at the listener's ear, to a long-term level of approximately 90 dBA.

The tracking task was similar to one devised by Gibbs in which the subject has control (using a joystick) of a vertical line on a screen with which he covers an illuminated point. Two hundred ms after he has covered the point, it reappears in another of five possible horizontally displaced positions, and the subject is again required to cover this point with the line. This process continued for as long as the time taken to process the 50 digit list.

Subjects

The subjects were 8 members of the staff of the IAM, aged between 20 and 28 years with normal hearing and vision. They were split into two groups of four, each group containing two males and two females. All subjects were instructed, trained and tested individually.

Procedure

Each subject was given a standardised introduction to the tasks and instruction on their performance. They were asked to perform as quickly and accurately as possible on each task, though paying primary attention to the maintenance of tracking task performance. They were then given six practice trials of combined listening and tracking; on the following day they were given another similar practice session. At the end of the second session the subjects had overcome the initial large practice effect on the tracking task.

Subjects were asked to attend on four test days. On two of these days they performed both tracking and listening tasks, and on the other two days only the listening task. Six trials were given on each test day, the first two being regarded as "warm-up" trials, the results of which were discarded. The remaining four trials consisted of one trial at each of the four noise levels, which were presented in a suitably balanced order. Each trial lasted as long as it took to complete the 50 digit list, and between trials the subject was allowed a 50 sec rest. Before each trial a 10 sec warning was given to the subject.

The measures recorded were:

1. Time taken to process the 50 random digit list.
2. The number of incorrect responses made in each list.
3. Hit rate on the tracking task.

RESULTS

The data from each group were subjected to a four factor analysis of variance, main factors being subject, noise levels, task type, and replicates. The analysis was split for noise and all the interactions of these factors were partitioned and tested.

The effect of noise on speed of list processing was highly significant ($p < 0.001$). Figure 2 illustrates how increasing the ambient noise level affects the time taken for list processing both with and without a concurrent tracking task. List processing error scores for all subjects were computed, but remained fairly constant and almost negligible at 2%.

Table 1 illustrates the effect of noise on tracking performance and it can be seen from this figure that tracking performance was maintained across all noise levels and did not change significantly.

Table 1. The Effect of Background Noise Level on Tracking Task Performance

Background Noise Level dBA	64	72	77	85
Mean Hit Rate	36.4	36	36.5	35.6

The magnitude of the change in communications performance induced by the addition of the tracking task appears large in Fig 2 but fails to achieve statistical significance.

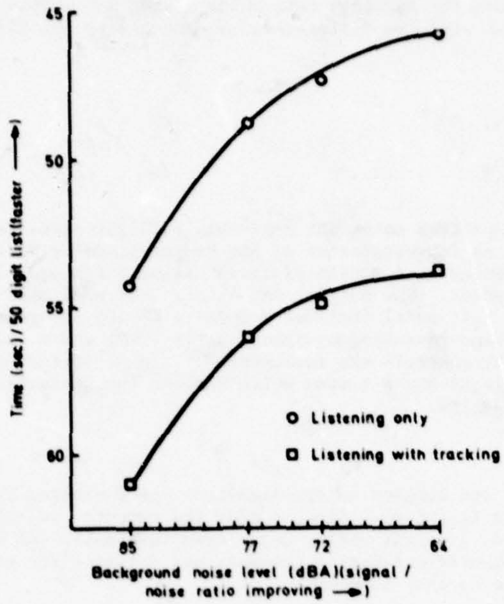


Fig 2 Effect of Background Noise and Tracking on Response Times (All Subjects Mean Values)

DISCUSSION

This experiment had a number of distinct aims. The first was to investigate whether the digit task developed by Holloway would be sensitive to the type of noise problem encountered in military aircraft where the problem is one of speech masking by broad band noise as opposed to the quantised distortion employed by Holloway. Curve A in Fig 2 shows clearly that performance on the test is markedly affected by the ambient noise level. Over the range tested, the mean total response time increased by about 20% with a 20 dB increase in background noise. This is a remarkable increase when one considers that a large part of the response time is taken up with the process of adding three to the number. Holloway has shown that the calculation time is independent of the noise level, so the actual increase in the time taken to analyse the incoming signal, and to decide what it is, must be very large.

The second object of the experiment was to see whether this communications task represented one which could be useful in investigating the relationship between the communications task and workload imposed by other tasks. There are reasons for supposing that this may be so other than the one mentioned in the introduction (ie that the task does not require a written response or a spoken response of which the accuracy is crucial). All normal intelligibility tests are presented at a rate of about one word every three seconds and even if spoken responses were accepted the rate could not be much increased. There is thus ample opportunity for the subject to "time-share" between the auditory and other tasks. The test used in this experiment provides the subject with no opportunity for any rest, however short, unless deliberately introduced by the subject, and it would in any case affect the results. The same can be said of the tracking task which, like the auditory task, requires continuous activity. Another possible advantage of this method was that although the subject is required to make only a spoken response, the accuracy of this response is not the statistic of importance as it is in normal intelligibility tests. In fact, if digits are used in the test, the accuracy is extremely high and errors can consequently be disregarded, but the reaction time remains a sensitive measure.

It is, however, also true that any real auditory task does not consist solely of identifying words, but also of comprehending them and utilising this comprehension in the formation of a suitable response. The normal intelligibility test does not, and is not designed to, involve these latter processes. The cumulative response time measure used in these experiments however, does necessitate some utilisation of the stimulus and can thus be argued to be more representative of the real-life situation.

Whether there exists an interaction between a tracking task and a communications task is a question which is non-trivial. The two tasks in this experiment use distinct sensory modalities in the presentation of information to the subject, and completely distinct response modalities. Hence it is only the extent to which each task requires access to some form of central processor which should determine the interaction. Incidentally, this independence of stimulus and response modalities represents a necessary condition for the success of any secondary task study - but one which is often not fulfilled in the literature.

Though this experiment established that the techniques developed were viable, it failed to establish a convincing effect of psychomotor workload on auditory processing. Experiment II was therefore devised

to pursue this problem under conditions where the workload was both more realistic and in some respects more controlled. In this experiment an attempt was made to clarify the relationship between the auditory task and a realistic flying task - in fact, flying an HS 125 aircraft simulator. Furthermore, the aircrew subjects were required to perform the auditory task while flying manoeuvres of varying difficulty in order to see whether the auditory task would be differentially affected by the difficulty of the primary (flying) task.

METHOD

Apparatus

Simulator

This investigation was undertaken using the Institute's flight simulator research facility. The device, a procedures trainer, is representative of the Hawker Siddeley HS 125, a two crew twin jet aircraft. The cockpit unit is mounted on a motion platform having pitch and roll axes. There is no representation of the external visual world. The subject sat in the left hand seat where he was provided with a full flight instrument panel. This panel includes a Collins FD 108 integrated flight system, an omni bearing selector (OBS), a distance measuring equipment meter (DME) and a radio magnetic indicator (RMI). All primary and secondary flight controls are represented. Simulated turbulence at low level was present throughout the task. The simulator has a system which enabled the controller or pilot to 'freeze' the whole system in any flight situation.

Computer

The flying performance of the subject in the simulator was monitored by a Digital Equipment Corporation PDP 12 computer. During a flight monitoring period the computer recorded, at one second intervals, the simulator airspeed, altitude and bank angle, and stored this data. At the end of each monitoring period of 30-50 seconds the computer calculated the mean and variance for each parameter and also the number of scores which contributed to that sample.

Communications Equipment

The communications for the flight procedure and for the intelligibility task were conducted through the Redifon Intercommunication system of the simulator. Both the experimenters and the subject wore Astrolite S21 headsets with standard magnetic microphones. White noise was injected into the "pacer"-to-subject link of the system from a white noise generator. The noise level was standardised by measuring the sound pressure level at the ear of the pilot when both simulator engines were operating at 75% throttle and adding sufficient white noise to bring this level to 87 dBA. Sound pressure levels were determined using a Bruel and Kjaer artificial ear and a Type 2603 microphone amplifier. The 87 dBA background noise level was selected as being a realistic operational cabin noise and also a level where complete recognition accuracy for the digits could be relied upon.

A tape recorder was used to record all conversation during the flight procedure.

The experimenter was also provided with a function box enabling him to control the flight monitoring periods of the computer, and equipment with which to record the cumulative response time in the communications task.

Tasks

Primary Task - Flying

Experiments were conducted at three phases of simulated flight which were considered to represent three levels of workload: namely

1. Sitting on the flight deck with the simulator controls "frozen" - baseline workload.

2. Flying straight and level at 1,000 feet and 210 knots using only the six primary flight instruments - low workload.

3. Flying round an arc, in a low bank turn, maintaining a distance of 6 n miles from the DME beacon at the centre of the arc, and maintaining the correct heading using the RMI system. Again, the simulator was to be flown at 1,000 feet and 210 knots - high workload.

Each subject flew two circuits with baseline measurements taken before and after the circuits and in the inter-circuit interval. A circuit consisted of a high workload sector and a low workload sector, these being presented in balanced design.

The experiment was also balanced for direction of turn, so that each subject flew round the arc in both clockwise and anticlockwise directions.

Secondary Task - Communications

The communications task was essentially the same as the one used in Experiment I except for the changes in intercommunication system noted above and for the use of only 30 random digits rather than 50.

A constant background noise level of 87 dBA was used in this experiment.

Subjects

Subjects were 8 members of the Experimental Flying Department RAE Farnborough aged between 28 and 49 years, all of whom had normal hearing and vision.

Previous aircraft and simulator experience was noted, and the time elapsed since the last flight.

All subjects were instructed, training and tested individually.

Procedure

Each subject was given a standardised introduction to each task and its performance. The need to maintain an accurate performance on the primary task (flying) was emphasised, and subjects were instructed to complete the secondary task (communications) as quickly and accurately as possible. Subjects were asked immediately after the experiment what strategy they had employed to comply with this requirement. After a brief explanation of the communications task subjects were given three practice trials of 30 digits.

Practice for the flying procedure was adapted for each subject according to his previous flying experience, but all subjects underwent a standardised procedure in the period immediately before the final briefing. This practice procedure involved flying in an arc centred on the DME radar beacon and 10 miles distant from it. It was thus similar to the "high workload" test procedure and gave the subject experience in the use of the navigation aid.

The instruction and practice lasted approximately one hour for each subject. After this, subjects were given a final briefing on the experimental flight procedure, explaining the order in which the high- and low-workload sequences were to be undertaken. This briefing was performed by the one experimenter and in the absence of the other who acted as "pacer". In this way the "pacer" could control the communications task with no knowledge of the workload conditions under which the pilot was operating and was therefore unable to bias the subject's performance.

The flight procedure comprised two circuits with a 5 minute rest interval. Each circuit comprised two sectors: a straight and level flight (low workload) and an arc six miles distant from the DME radar beacon (high workload). During each of these sectors two 30 sec samples of flying performance were monitored by the PDP 12 computer. There were also two periods where the communications task ran concurrently with the flying task; these lasted for 30-60 seconds, the time being recorded as the subject's cumulative response time. Flight performance was also monitored during these periods. At the beginning and end of each circuit the simulator controls were "frozen" and the communications task was run once to provide baseline data.

The data collected were, therefore, an assessment of the primary (flying) task performance in terms of altitude, airspeed and bank angle, and performance on the communications (secondary) task in terms of cumulative response time and accuracy of processing the 30 digits.

There was also a questionnaire administered after the experiment concerning the accuracy with which real-life tasks were simulated and the strategies which the subjects employed.

RESULTS

The cumulative response times from the communications task were subjected to a split-plot four-factor analysis of variance; main factors being subjects, workloads, circuits and replicates. All interactions of these factors were partitioned and tested.

The primary task workload was found to have a highly significant effect on the speed of digit processing ($p < 0.01$) such that the greater the workload in the primary (flying) task, the slower was performance of the secondary (communications) task (Fig 3). Accuracy in the communications task was found to be consistently high (> 97%) for all subjects regardless of primary task workload.

Individual differences between subjects was also a significant factor but there were no other significant main factors or interactions.

An analysis of the flight data for each of the parameters airspeed, altitude, bank-angle, was carried out to see whether the introduction of a secondary task had any effect on the primary task performance. This analysis was a five factor analysis of variance with main factors being subjects, workload, circuits, replicates and the presence or absence of the concurrent secondary task.

The analysis showed significant individual differences between subjects for airspeed and altitude ($p < 0.1$) and ($p < 0.01$) and a significant effect of workload on altitude ($p < 0.01$). Other main factors and interactions were non significant. There was no effect of the secondary communications task on primary (flying) task performance.

DISCUSSION

The experiments described here demonstrate that a speech communications task which uses response times as a measure is sensitive both to signal quality and concurrent workload. It should not be construed, however, that precisely the same auditory and cognitive processes are affected by degradation of signal quality and by workload. The simplest approach to the activity of hearing and replying to any acoustic signal must include the processes of acoustic and semantic analysis, comprehension and decision making, and response organisation. It is probable that in the first experiment the process of acoustic analysis was slowed by a reduction in the number of cues available in the poorer signal. Holloway's work strongly suggests that this is so, as he found that for the type of distortion in which he was interested produced a constant increase in total response time irrespective of the complexity of the manipulation required of the subject. That is to say, it did not matter whether the subject was required to multiply by two, add

three, or perform any similar calculation - introducing distortion still produced a constant increase in list processing time.

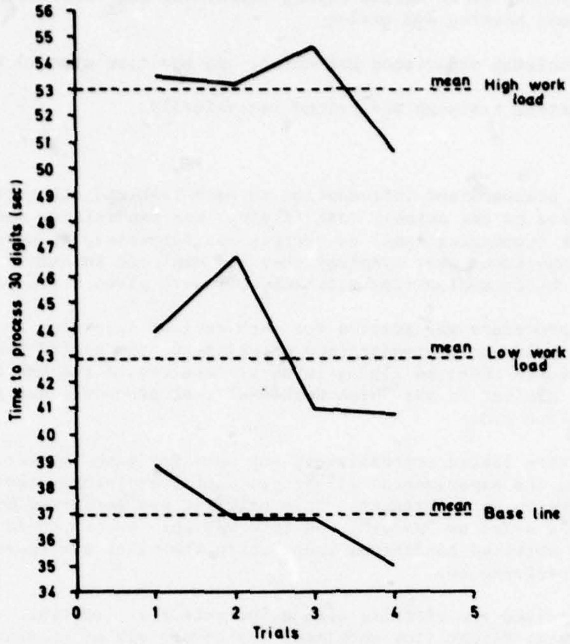


Fig 3 Effect of Cockpit Workload on a Auditory Task

In the second experiment it has not been demonstrated whether the workload had its effect by slowing the perceptual processes, by making less capacity available for the process of adding three to the perceived number, or even by reducing access to the mechanism which selects and organises the motor activity of the spoken response. It would be of interest to devise experiments to test which of these processes is affected, but this is not necessary in the present context as all are involved in responding to normal conversational speech.

The particular problems addressed by this paper have been little studied, but the work of Broadbent (1958) merits discussion. In his experiment a tracking task was shown to reflect difficulty changes in simultaneous performance of an auditory task induced by various different frequency shifts. The particularly interesting point about this experiment was that performance on the tracking task was able to discriminate between speech conditions of apparently equal intelligibility, but of different subjective listening difficulty.

It therefore appears that there are at least two attributes of a speech signal of human factors importance. The first of these is the absolute quantity of information carried in the signal (which will define the intelligibility of the signal) and the second is the amount of processing required of the listener to recover the amount of information (which will define the degree of subjective difficulty encountered in listening to the signal). As Broadbent used a standard intelligibility test to evaluate the amount of information present in the signal, the tracking task was used, successfully, to indicate the amount of processing involved in the intelligibility test.

It can be argued that the response time measure used in the present experiments is itself an index of the amount of processing required to gain the available information from the signal (assuming that more processing requires a longer time). Thus this task can be said to give a better indication of the overall human factors demands of a given signal than can a conventional intelligibility test, unless the conventional test is paired with a tracking task as in the experiment of Broadbent. However, the difficulty of recovering good intelligibility data in such experiments is larger as spoken responses must be used with the inherent marking difficulties of using this response mode.

The simulator task in the present experiment therefore was not playing the same role as the tracking task in Broadbent's experiment. Broadbent was interested in using tracking performance to indicate the workload imposed by the auditory task. In the present experiment the simulator imposed a realistic psychomotor workload in an attempt to see to what extent this interfered with a concurrent auditory workload. Experiments by Johnston (1975) also attempted to investigate psychomotor workload (a tracking task) as a source of interference with an auditory task, but unfortunately the results of these experiments are inconclusive due to poor design and analysis. Other experiments suffer similar limitations.

A further aspect of the experiment reported here is that no deleterious effect of the auditory task was found on the primary tasks of tracking and simulator flying on which the subjects had been told to attempt to maintain their performance. This result could have occurred if the measures of flying performance used were so insensitive that they were not able to register real changes in flying performance. It

is argued that this is not the case, as the altitude measure showed changes dependent on primary task workload, but not of the presence or absence of secondary task. This is interesting in view of the fact that Rolfe (1969) describes the major difficulties associated with the use of secondary tasks as being interference with primary task performance. Poulton (1975) comprehensively reviews recent work in which tracking tasks were performed with another task, and this review makes clear the difficulty of not interfering with the primary task. If, as seems likely, the auditory task described here can be used in a way which does not interfere with primary task performance, it is of possible value as a measure of primary task workload in situations where the experimenter would not normally be concerned with auditory performance.

It appears then, that while this task was devised to investigate the way in which psychomotor performance and communications performance interact, it may be useful as a traditional form of secondary task - but a secondary task which appears relatively free from the traditional objections to secondary tasks.

Perhaps the most important recent paper to discuss the theoretical implications of the use of secondary tasks is that of Norman and Bobrow (1975). In this paper, the problem of how the quality of the incoming signal (the data) and the amount of available mental capacity (resources) affect the final performance of the human operator is discussed.

In the present experiments it would be of interest to know which aspects of the overall task of responding to digits are affected by limited data and which by limited resources.

Figure 2 shows that a 50 digit test may be processed in, for example, 54 secs in two ways:

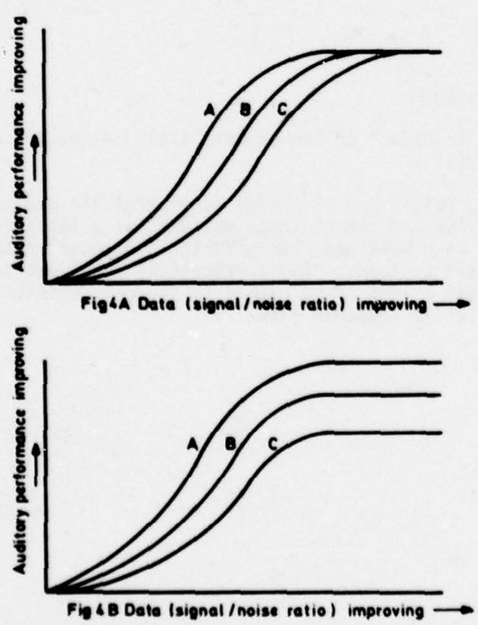
1. By listening to the list in quiet (65 dBA) but while tracking.

2. By listening to the list in a high noise (85 dBA) environment but not while tracking.

It thus appears that there is a trade-off between available resources (or "spare mental capacity") and data (in this example, signal/noise ratio). It would be wrong to suggest however, that a decrement caused by increased workload may always be compensated by improving the signal/noise ratio, for the reason outlined below.

It is well known that the way intelligibility varies with S/N ratio is always described by some sigmoid function. Thus the curves in Fig 2 may be regarded as sections of larger sigmoids. However, as the curves in Fig 2 are only part of a larger family of curves, we have not sufficient information to decide whether the family of curves will appear as in Figs 4a or 4b. Either of these families of curves are consistent with the results of the experiments described.

The differences between the families are that if the curves are as in Fig 4a a given workload may always be compensated by improving signal quality. In contrast however, the curves depicted in Fig 4b become "resource-limited". This means that although over a certain section, workload and signal quality may be traded-off one with the other the presence of a resource limitation (limited spare mental capacity) produces a limit to the performance possible, no matter how much signal quality is improved. It is suggested that though the evidence is extremely weak, Fig 2 would fit better into the model described by Fig 4b than that described by Fig 4a.



A High resource condition (large spare capacity)
B Medium resource condition (moderate spare capacity)
C Low resource condition (low spare capacity)

Fig 4 Possible Interactions of Data and Resource in an Auditory Processing Task

It is also true, as mentioned earlier, that even the relatively simple task of responding to a digit involves a number of distinct processes (acoustic and sematic analysis of the signal, performing the arithmetic manipulation, organising and making a response) and from the present experiment we have no way of knowing whether resource and data limitations reduce the overall level of performance by affecting the same or different parts of this chain of processes.

Knowing the locus of the interference produced by a psychomotor task on an auditory processing task is important as it is known that aircrew complain of fatigue which they attribute to poor communications quality. An experiment such as experiment II will indicate that they *do* have a genuine case in that it is clearly demonstrated that the task of auditory processing is likely to increase considerably their overall workload. However it may be misleading to suggest that the cure would be to improve the quality of the auditory signal, as in addition to the limitations described above, doing this may merely improve performance at one locus in the processing chain (for example the acoustic stage of analysis) and in this way compensate for reduced performance at the real locus of interaction between the tasks (which may be the semantic analysis stage - understanding the signal). Thus, in the limiting case, the solution would not necessarily be to improve the acoustic quality of the signal, but to make it simpler to understand.

It may be concluded that the digit processing task described above is likely to be useful in the study of the interaction of auditory processing and other workload, as the experiments described have shown the task to be sensitive both to the effects of changes in signal quality and to the level of primary task workload. This may enable some of the theoretical problems associated with the interaction of auditory processing and workload which are discussed above to be resolved. In addition, the task appears to interfere minimally with primary task performance and may thus be useful as a secondary task in the assessment of primary task workload.

REFERENCES

1. BROADBENT,D.E. Perception and Communication. Pergamon Press. 1958.

2. GIBBS,C.B. Probability Learning in Step-Input Tracking. J. Psychol. 56. 1965. pp. 233-242.

3. HOLLOWAY,C.M. Some Effects of Noise on a Speech Communication Task. Sound 6 pp. 27-31. 1972.

4. JOHNSTON,M.E. The effect of a tracking task on speech intelligibility in noise. RAE Technical Report 75014. 1975.

5. NORMAN,D.A. and BOBROW,D.G. On Data Limited and Resource Limited Processes. Cognitive Psychology 7. 1975. pp. 44-64.

6. POLLACK,I. Reaction Times to Unknown Word Sets in Noise. Language and Speech 6 pp. 189-195. 1963.

7. POULTON,C. Tracking Skill and Manual Control. Academic Press. 1975.

8. ROLFE,J.M. The Secondary Task as a Performance Measure: A Survey of the Literature. RAF Institute of Aviation Medicine Report No 474. 1969.

DISCUSSION

SANDERS: (Netherlands) — Does the secondary task have an effect on the primary task? Also, how did you measure "digit processing time?"

GREEN: (United Kingdom) — Performance parameters (e.g., height and airspeed) were carefully measured on the primary task of simulator flying and though these measures were sensitive in themselves to changes in primary task workload, the effect of introducing the secondary task could not be detected on them. "Digit Processing Time" was taken to be the total time taken to process a list of 30 digits in the way described in the paper: it is a form of cumulative response time.

PITCH AND FORMANT ANALYSIS OF THE VOICE IN THE INVESTIGATION OF PILOT WORKLOAD

by

Cannings, R., Borland, R.G., Hill, L.E. & Nicholson, A.N.
Royal Air Force Institute of Aviation Medicine
Farnborough, Hampshire, United Kingdom

INTRODUCTION

Assessment of pilot performance in the dynamic flight situation is a complex and difficult problem, but a useful approach is to study the ease with which the task is being executed, and so the reserve available to the pilot to handle additional workload. Understanding the capability, state of preparedness or potential of the pilot necessitates information on the manner in which the nervous system is being modified by the flight situation, and on the significance of modified central nervous activity to the performance of a flying task. It is in this context that we have carried out studies on heart rate and finger tremor during the short duration high workload of the let-down, approach and landing[1,2], and have attempted to correlate these observations with pilot assessments of workload based on the part played by the many factors influencing the complexity and difficulty of the flying task[3].

In the previous studies[2] it was observed that during an uneventful let-down there was a decrease in rr interval and increase in finger tremor, but that during let-downs in which poor control was accompanied by inadequate aids and unfavourable meteorological conditions and frequently preceded by a high workload cruise, a much greater decrease in the mean rr interval around touch down was observed, though the increase in finger tremor was not exaggerated. On the other hand, in the event of an unresolved problem persisting or a fresh problem of some magnitude appearing during the approach low mean rr intervals were often accompanied by gross finger tremor, and it was suggested that the mean rr interval around touch down reflected overall workload whereas unusual increases in finger tremor were associated with untoward events.

In further experiments[4] the workload of the pilot during the approach was modified by coupling the aircraft to the ILS localiser and glide slope path (coupled approach), or by increasing the participation of the co-pilot in the handling of the aircraft (shared approach). The mean rr intervals around touch down of the coupled approaches were increased compared with let-downs of equal difficulty handled throughout by the pilot (manual let-down). In the shared approaches the relationship between the mean rr interval around touch down and workload was similar to that for manual let-downs, unless the co-pilot handled the aircraft to 500 feet when the mean rr interval around touch down was increased over let-downs for a wide range of difficulty. The appearance of finger tremor was not affected by the modified workload pattern. It was concluded that workload during the initial part of the approach may influence the neurological state of the pilot around touch down, but that finger tremor remained an indicant of untoward events.

In the studies of subjective assessments of workload[2] the pilot's assessments of the individual factors of workload i.e. aircraft serviceability, aids, control, meteorology etc. were related to overall workload. The accuracy of predicting overall workload from assessments of individual factors was greater during let-downs of limited difficulty, and the model became more complex as the individual factors became more unfavourable. A linear model became less satisfactory as the individual factors influencing workload became less favourable and these observations raised two issues. A pilot may have used a different process to assess overall workload under difficult circumstances, or the stress which accompanies high workload may have impaired his processes of computation bringing a greater variability to his assessments.

In view of these studies it was considered that changes in central nervous activity and subjective assessments of workload may prove to be useful techniques in the investigation of workload in aircrew. However, such studies require instrumentation of the pilot, and not inconsiderable attention to the problems of subjective assessment, and so we have turned our attention to the analysis of the voice which may provide information on the activity of the central nervous system.

Stress and High Workload in the Speech Signal

Many studies have been performed to evaluate the effects of stress on human performance, but few have considered the spoken word as an index of central nervous arousal. In part, this has been due to the difficulties in establishing known and repeatable stressful conditions, but the quantitative analysis of the speech waveform has also presented a problem. In recent years digital computers have stimulated much interest in automatic speech recognition and synthesis, and, as mathematical techniques have been developed to provide a precise description of the speech waveform, voice analysis may prove a useful technique. Stress may be defined as task or failure induced[5] with motivation playing a role in the interpretation of each type. In the flight deck, task induced stress is a function of workload (both physiological and psychological), while failure induced stress will only result under catastrophic circumstances. Our investigations use workload stress, though other studies have involved the latter situation[6,7].

The techniques used in the analysis of the voice have included manual or semi-automatic analyses of spectrograms derived from speech communications in fatal or near fatal accidents, and may be broadly classified into two groups. The first uses a wideband analysing filter (200-400 Hz bandwidth) which tends to emphasize resonances within the voice spectrum, while the second uses a narrow band analysing filter (<50 Hz bandwidth) and highlights harmonic structure. Though increases in fundamental pitch have been demonstrated, more subtle variations in voice characteristics are to be expected during a routine let-down approach and landing. Indeed, spectrographic pitch determination during laboratory produced stress conditions has shown inconsistent inter-subject variability[8]. This, together with the fact that additional information may only be derived from the spectrogram by subjective means, suggests that a more detailed analysis may be appropriate. In this context voice data during space flight[9] has been analysed by analogue techniques, and the results appear to be promising.

In view of these considerations the approach adopted for the present study was to use computer techniques to extract parameters from the speech waveform which would be amenable to statistical analysis, and to compare the structure of a pilot's speech waveform at different points in the flight profile. Physiological data and subjective assessments may be used to indicate the workload of the flight profile, and so it should be possible to correlate voice parameters with known workload[2]. In the present study we concentrated on the call sign of British Airways - "Speedbird", and by using a single work both pitch and formant information has been analysed.

Basic Speech Waveform

It is convenient to consider speech as having two structural aspects[10]. The speech waveform may be split into speech and pause segments, and the speech segments may be further divided into voiced an unvoiced intervals. Voiced intervals represent vocal chord activity producing glottal pulses at the fundamental pitch frequency, and are rich in harmonics. These pulses may be modified by resonating cavities (pharyngeal, oral and nasal) within the vocal tract. Such activity produces vowel and vowel-like sounds. Unvoiced intervals are characterised by fricative sounds produced by continuous turbulent airflow through constrictions in the vocal tract, i.e. articulators such as the lips, teeth and tongue. These have the effect of wideband noise sources which may be produced in conjunction with vocal chord activity. However, pitch determination is difficult under these conditions due to the masking effect of the noise.

It is sometimes necessary to consider speech as having a third structural segment which is a transitional phase between the speech and pause intervals. This phase which may be voiced or unvoiced produces the so called stop or plosive sounds, but it is considered unnecessary to include it in a stress analysis. English, in common with other western languages, is non-tonal in nature, and during voiced speech segments information exchange between speaker and listener does not require changes in fundamental pitch, but rather changes in vocal tract resonances. It would be expected that pitch level remains essentially constant during an utterance, so that any long term variations may be a primary or secondary function of the emotional state of the speaker. For example, a stressful condition may directly affect the motor output to the vocal chords, while, on the other hand, it is known that the respiratory pattern is disturbed by stress, and this in turn may alter fundamental pitch via changes in sub-glottal pressure. Pitch is basic to the spoken word, and may be chosen as one of the most important speech parameters.

The resonances produced in the vocal tract are known as formants and comprise further important parameters of the speech waveform. Their variations are characteristic of the information content of an utternace and no repeatable patterns are to be expected in random phrases drawn from the speech waveform. However, formant profiles during the repeated utterance of the same word should be similar, and any gross variations will not be a function of the information in the speech signal. Changes in the formant profile of a given word may be indicative of stress.

Finally, the low frequency envelope around the speech waveform may be used to characterise a continuous section of speech. In particular, word length, speech pause onset and transitional phases as well as relative loudness levels may be assessed. While this is a useful approach, the discontinuous nature of cockpit communications presents a major problem, and the present study is restricted to pitch and formant analysis, though attempts are being made to develop methods which will allow statistical analysis of the waveform envelope.

Analysis Techniques

Before computer techniques can be used to analysed the voice, a model must be proposed which simulates the action of the vocal apparatus to a required degree of accuracy. The model which is described here has been used in both analysis and synthesis studies. Reconstruction of the speech waveform is a useful method of testing the validity of the model, and though in the present case only a simulation of voiced speech waveforms is proposed, the model does appear to behave realistically under a broader class of conditions, including fricatives[11].

Fig. 1. illustrates the model in which the vocal pulse generator produces air impulses at the fundamental pitch frequency. The impulse train whose size is determined by a gain constant is modified by a linear filter, $G(s)$, representing the effect of the glottal source. The vocal tract resonators are represented by a cascade of linear quasi-time invarient filters $H_i(s)$. In general four formant filters are sufficient to describe the vocal tract. The structure of each filter is shown in Fig. 2.

Thus
$$H_i(s) = \frac{1}{s^2 L_i C_i + s R_i C_i + 1}$$

so that
$$\zeta_i = \frac{R_i}{2}\sqrt{\left(\frac{L_i}{C_i}\right)} \quad \text{and} \quad \omega_i = \sqrt{\left(\frac{1}{L_i C_i}\right)}$$

It therefore follows that the spectral peak of a particular formant occurs at

$$\omega_{imax} = \sqrt{\left(\frac{1}{L_i C_i} - \frac{R_i^2}{2L_i^2}\right)}$$

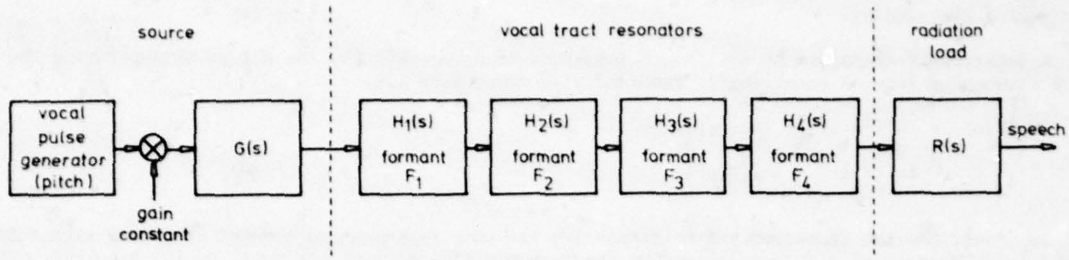


Figure 1

Simplified model of voiced speech production

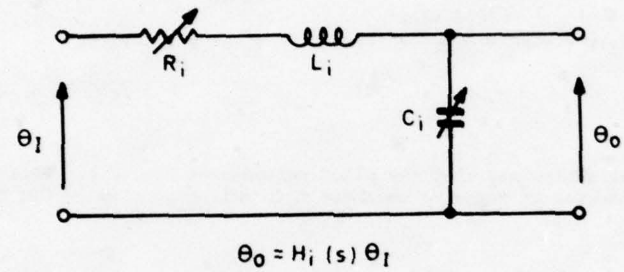


Figure 2

Electrical analogue of a single formant resonator

Physically, L, C and R summarises the acoustic properties of air motion in a cylindrical tube. Thus, L is an acoustic intertance, C is a compliance term which is dependent upon the cross-sectional area of the vocal tract, and R is a viscose drag term, dependent upon both cross-sectional area and circumference of the vocal tract. Formant frequencies may be varied by means of changes in the capacitance C_i while formant bandwidth is controlled by R_i. Although the latter parameter is of secondary importance, both vary relatively slowly (in phase with the vowel content of an utterance) and so the formant system may be regarded as invariant, at least in terms of short-time analysis. In this context the process is considered stationary during periods of 100 ms of less. A final filter R(s) represents the radiation load which physically may be interpreted as the loading effect produced as sound leaves the vocal tract and enters the external atmosphere.

The effects of G and R are normally lumped together since both are considered as time invariant. For a given speaker G(s)·R(s) may be evaluated from a long time average spectrum (formant peaks will average out), but generally a spectrum of the form given in Fig. 3. is adequate.

Pitch and formant extraction using the model of Fig. 1. is best approached by considering the speech output signal f(t) to be a convolution of the vocal source impulse train with the impulse responses of the system filters. Thus

$$f(t) = p(t) * g(t) * h(t) * r(t) \qquad - (1)$$

where h(t) is the impulse response derived from the system

$$H(s) = \prod_{i=1}^{4} H_i(s) \qquad - (2)$$

and * denotes convolution. The problem resolves therefore into the separation of the four members of the right hand side of equation (1). One of the more efficient techniques for doing this is known as cepstral

analysis,[12,13,14]. If the speech output signal is further reduced to

$$f(t) = s(t) * h(t) \qquad - (3)$$

where s(t) represents the output of the vocal source (and may also include the effect of the radiation load) then frequency transformation of equation (3) yields

$$F(\omega) = S(\omega) \cdot H(\omega) \qquad - (4)$$

A logarithmic transform of the speech magnitude spectrum then has the effect of separating the members of the right hand side of equation (4) into additive components i.e.

$$\ln|F(\omega)| = \ln(|S(\omega)| \cdot |H(\omega)|)$$

$$\ln|S(\omega)| + \ln|H(\omega)| \qquad - (5)$$

$\ln|F(\omega)|$ has the appearance of an undulating function representing formant structure with a superimposed "high frequency" ripple representing the harmonic structure of the vocal source spectrum. The additive nature of equation (5) is maintained during inverse frequency transformation which results in the cepstrum (or inverse spectrum). The cepstrum has units of quefrency (or inverse frequency) which are measured in milliseconds. Clearly, the harmonic structure in the log magnitude spectrum manifests itself as a sharp peak in the cepstrum from which pitch period may be determined. Finally, short time filtering the cepstrum to remove the pitch peak and transforming back into the frequency domain removes the high frequency ripple from the log magnitude spectrum, thus enabling a peak picking technique to be implemented to extract formant frequencies. Short time filtering is effected by multiplying the cepstrum by

$$x(t) = \begin{cases} 1 & : \quad T_c - \delta t > t \\ 0.5\left(1 + \cos\left(\dfrac{\pi(t - T_c + \delta t)}{2\delta t}\right)\right) & : \quad T_c - \delta t < t < T_c + \delta t \\ 0 & : \quad t > T_c + \delta t \end{cases}$$

where T_c is chosen to be slightly less than the pitch period (see Fig. 4.). This process is analogous to low pass filtering a time series by removing unwanted spectral components in the frequency domain.

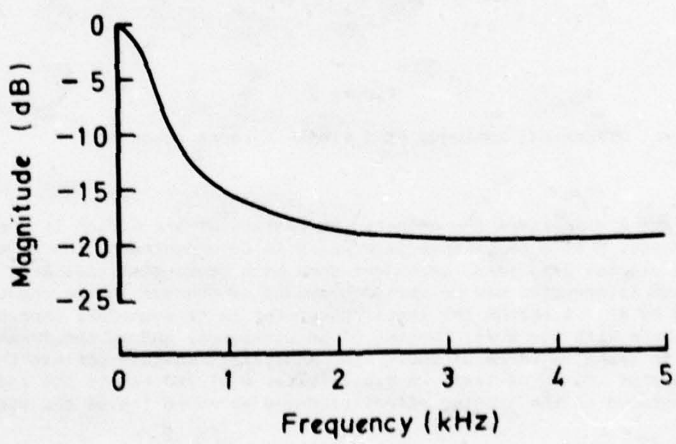


Figure 3

Magnitude spectrum of the function G(s) . R(s) describing the effect of the source spectrum and radiation load

It should be stressted that cepstral analysis is not the only method of extracting pitch and formants but apart from its efficiency, it is probably one of the easier methods to impliment. The autocorrelation function for example is defined in a similar fashion, but is less effective since it lacks the logarithmic transform which separates the source and vocal tract aspects of the voice spectrum into additive components. It should also be pointed out that there are several definitions of the cepstrum: the one used in this study defines it as the inverse Fourier transform of the log magnitude spectrum. Phase information in the cepstrum is disregarded, this being of no consequence since we are only concerned with frequency extraction from magnitude plots.

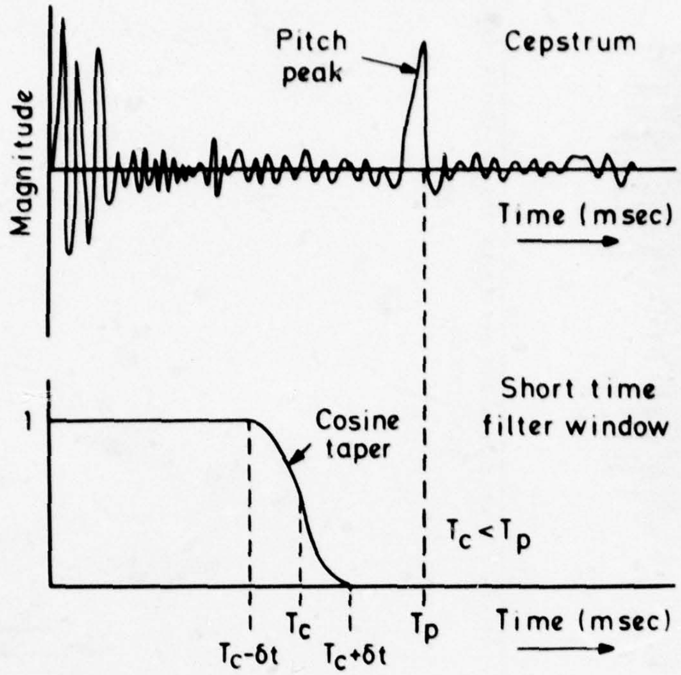


Figure 4

Short time filtering technique to remove the pitch peak from the cepstrum

Implementation of Cepstrum Analysis using Voice Data Recorded in a Cockpit Environment

We are concerned with the problem of using voice data recorded in real situations rather than in anechoic conditions. Clearly, attempts to correlate amplitude information between utterances must be disregarded since the exact orientation of the pilot with respect to the microphone boom is unknown and any attempts to limit relative changes in position would disturb the stressful nature of the situation. Furthermore, ambient noise levels in the cockpit, due mainly to aircraft vibration and engine noise introduce spurious signals into the voice data. Figure 5 illustrates a typical in-flight utterance of the word "Speedbird" (from touch down into Caracus, October 1974) and one recorded under near anechoic laboratory conditions. In-flight recordings were made using a high quality condenser microphone (Sennheiser) and B & K tape reocrder with 9 KHz bandwidth at 1.5 i.p.s. and 60 KHz bandwidth at 15 i.p.s. The time series data illustrates the poor signal to noise ratio in the field recordings, while comparison of the long time spectra suggests that the greatest part of the noise is concentrated below 150 Hz. Although this may be removed by a low pass filtering technique, it is likely that the pitch fundamental will also be removed. However, cepstral analysis is relatively insensitive to such an operation since pitch is extracted from the harmonic structure of the data. With reference to the long time power spectra it is of interest to note that pitch information is easily obtained from the laboratory data (1st, 3rd, 4th and 5th spectra peaks are equispaced while the 2nd peak probably represents the 1st formant) but again noise corruption does not allow this in the field data.

A schematic diagram of the complete analysis system is given in Fig. 6. Speech digitisation was carried out at 8192 Hz giving a maximum signal bandwidth of 4096 Hz. It was decided not to utilise information in higher order formants due to the inherent quality and bandwidth limitations of normal radio transmissions, which would ultimately be used. However, the seemingly high sampling rate was chosen to produce sufficient ripples in the log magnitude spectrum to ensure an adequate pitch peak in the cepstrum. Furthermore, Fig. 7. illustrates the effect of sampling rate upon frequency resolution in the cepstrum, assuming that Fast Fourier Transform methods are implemented. It is clear that the lower sampling rates introduce unacceptable quantisation errors as the calculated pitch frequency increases.

It is stressed that the computer software was designed to extract speech parameters from a noisy environment rahter that produce an exact analogue of the pilot's voice, and in this context experience has shown that techniques such as linear trend removal in the log magnitude spectrum have a desirable effect.

Initial processing was restricted to the vowel sound in the second syllable of "Speedbird". A 0.5 sec section of speech containing this sound was digitised with the aid of marker pulses placed on the original recordings. Visual inspection of the long time power spectrum was used to establish whether or not excessive noise components should be removed. If so, this was achieved by reconstructing the original time series, having filtered the unwanted noise peaks from the power spectrum. However, it was still necessary to intorduce introduce an adaptive threshold technique to define the required speech epoch. Due to the way in which

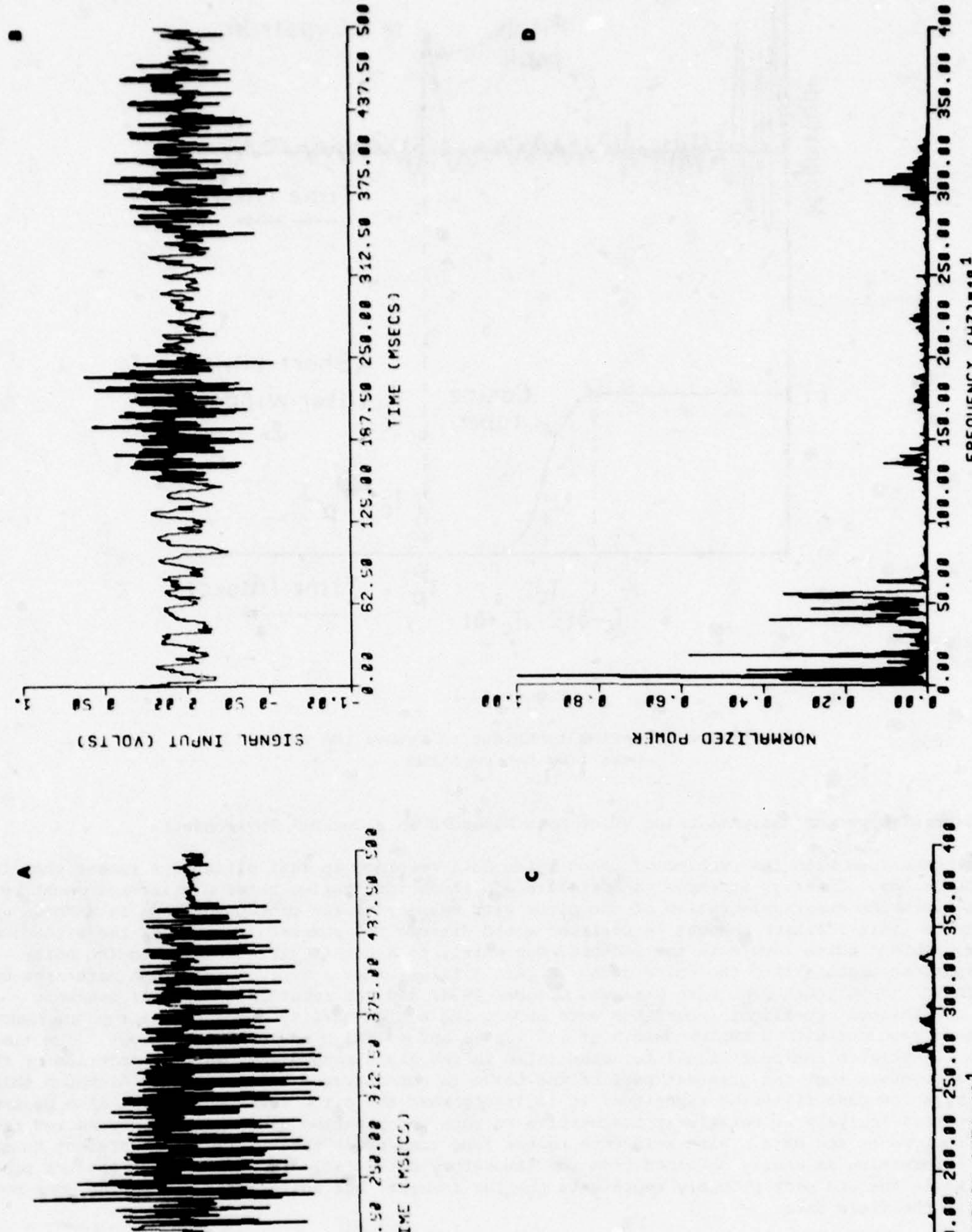


Figure 5

Time series and long time spectral plots of the word "Speedbird" recorded during near anechoic laboratory conditions (A and C) and during a landing into Caracus (B and D)

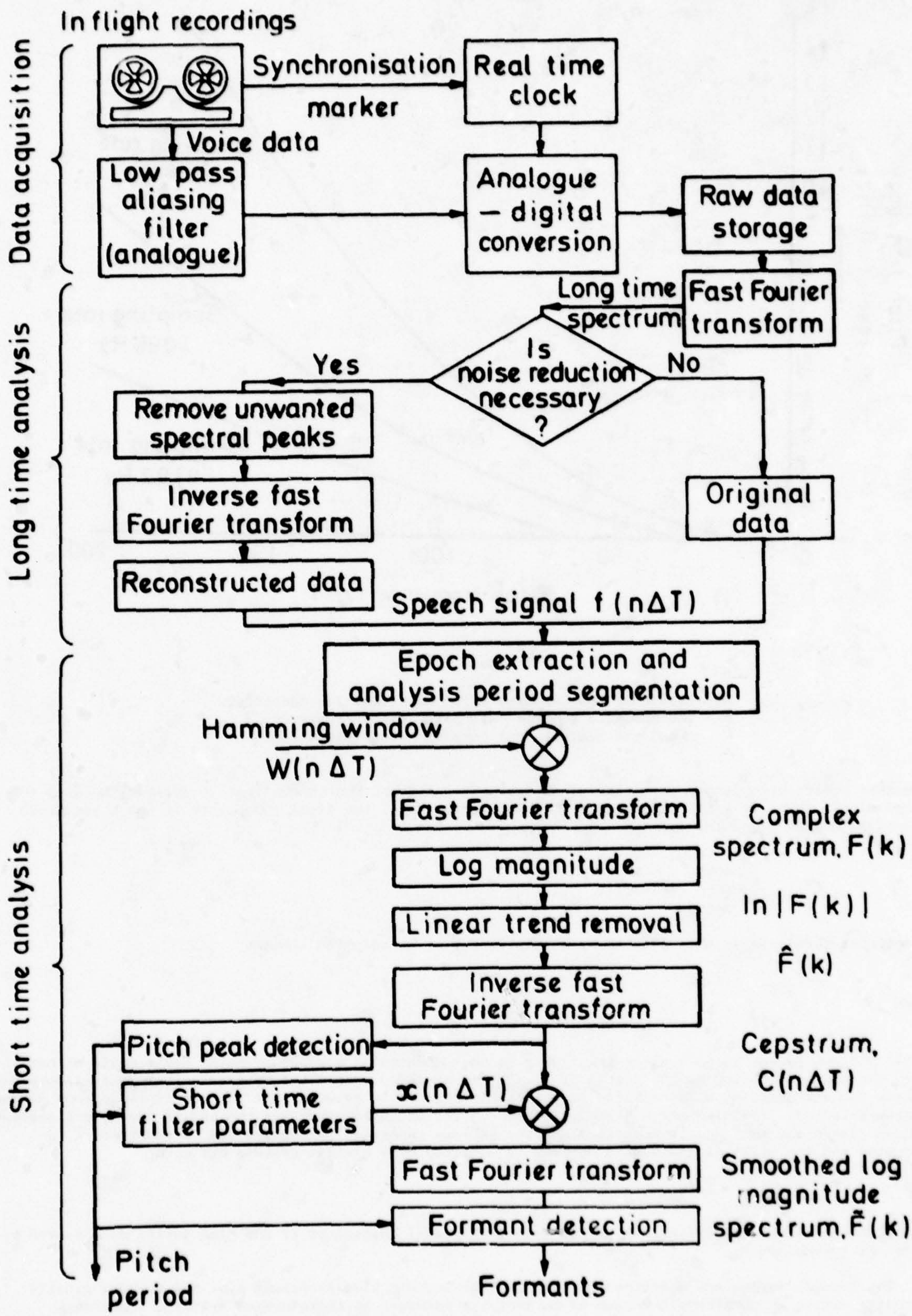


Figure 6

Schematic representation of the complete analysis system.
The "short time" analysis is performed on each 25 ms segment of data

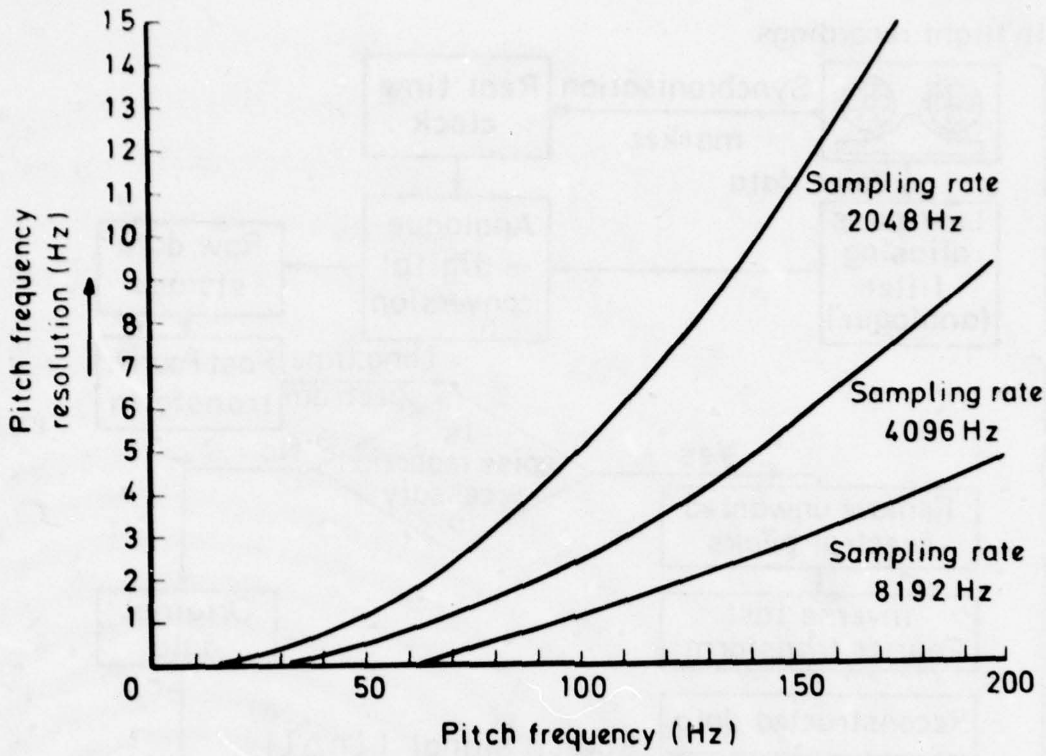


Figure 7

Resolution of the cepstral pitch determination technique.
The abscissa denotes calculated pitch frequency
and the ordinate the corresponding resolution

the marker pulses were placed on the original tape, it was found most convenient to analyse the data in reverse time. Thus, if $f(n\Delta T)$ represents the time series data, the final 100 points in the ½ sec epoch are used to define a modulus average

$$\bar{f} = \frac{1}{100} \sum_{n=3997}^{4096} |f(n\Delta T)|$$

An amplitude threshold is then defined by multiplying $\bar{f}$ by an adapting element

$\bar{f}_{ADPT}$ i.e.

$$f_{THRESH} = \bar{f} \cdot \bar{f}_{ADPT}$$

Initially $\bar{f}_{ADPT}$ is set to be greater than 1. $\bar{f}$ is then updated on a point by point basis until either the moving average has shifted by 100 points or f_{THRESH} is exceeded. In the former case f_{THRESH} is recomputed to allow for any drifting effects while in the latter case it is assumed that the end of the speech segment has been detected, providing $n\Delta T > ½$ sec. If $n\Delta T < ½$ sec it is assumed that the end of the speech segment has been missed and so $\bar{f}_{ADPT}$ is decreased and the process repeated. Otherwise $n\Delta T$ is recorded, $\bar{f}$ is recomputed and the start of the speech segment is detected by a similar process but with

$$f_{THRESH} = \bar{f}/\bar{f}_{ADPT}$$

This relatively simple algorithm was useful even when visual inspection of the time series data failed to define the speech epoch.

The formant resonators described by equation (2) are not time invariant i.e. the formant profile may change during an utterance. Choice of an analysis interval is therefore of critical importance. Ideally, it should be approximately equal to the period of the lowest formant, but this would not span a complete pitch period and so pitch determination would be less certain. A realistic compromise is to chose a period of 25 ms which covers about 3 pitch cycles, during which time formant changes are slight. However, each analysis interval is overlapped so that effectively pitch and formants are evaluated every 15 ms. Since a 25 ms analysis interval represents 205 points in the time series, it is necessary to pad the interval with zeros for transformation purposes. A resolution of 8Hz in the frequency spectrum is obtained with a 1024 point transform which should also be sufficient to prevent aliasing in the cepstrum.

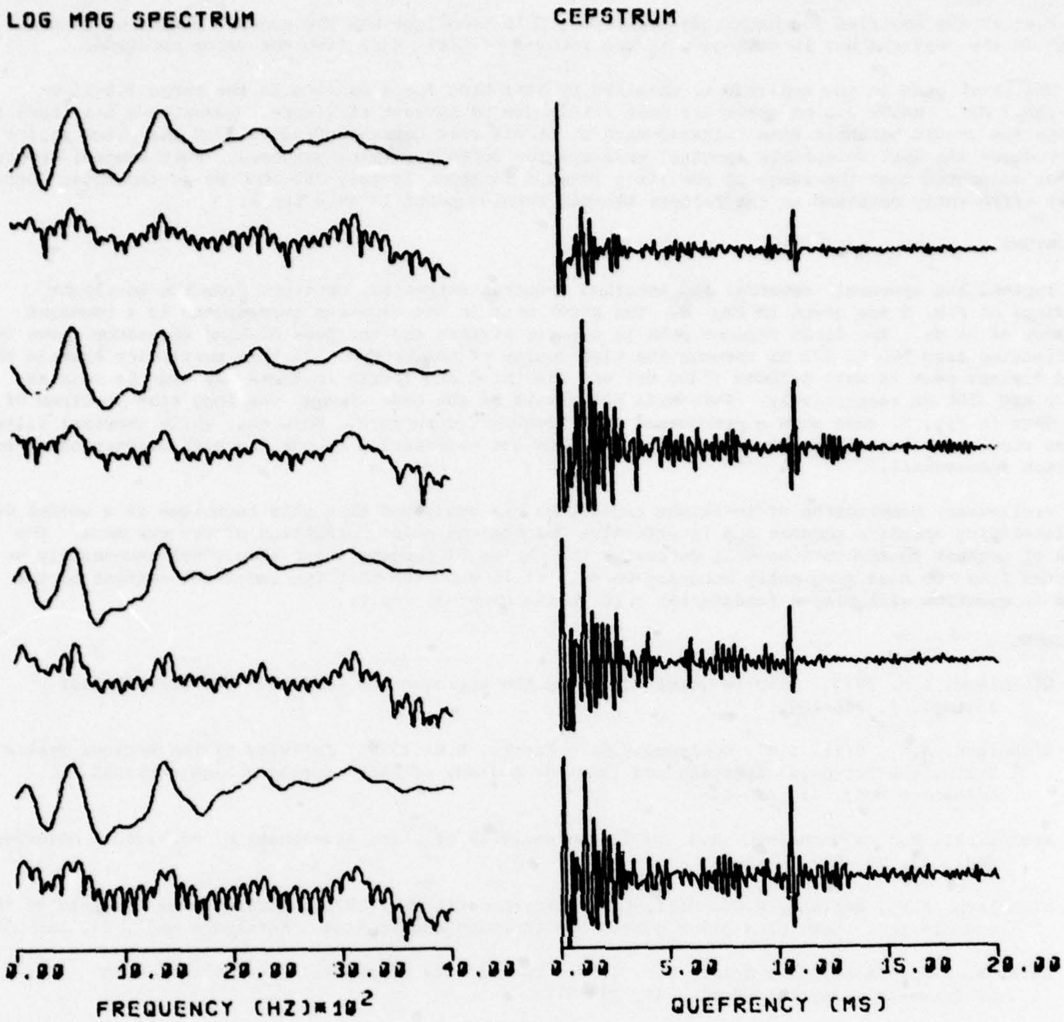


Figure 8

Spectral, cepstral and smoothed spectral data obtained from the in-flight recordings of Fig. 5. Each set of 3 curves represents a 25 ms analysis interval. The time course of the phonem progresses from the bottom to the top of the diagram

It is unlikely that the analysis interval is pitch synchronous and so a rectangular data window would have an adverse effect on the magnitude spectrum. The time series data was therefore multiplied by a Hamming window which minimises leakage providing the spectral components do not oscillate in resonance with the side lobes of the window. The Hamming window is given by

$$w(n\Delta T) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n\Delta T}{N\Delta T}\right) & : \quad 0 <= n\Delta T <= N\Delta T \\ 0 & : \quad \text{elsewhere} \end{cases}$$

Linear trend removal in the log magnitude spectrum was achieved by calculating a linear regression on the data points with mean removed, and modifying them accordingly. Thus

$$F'(k) = \ln |F(k)| - M, \qquad k = 1,N$$

where $$M = \frac{1}{N} \sum_{k=1}^{N} \ln |F(k)|$$

and evaluating $$C = 6 \sum_{k=1}^{N} \frac{(2k-N-1)\ \ F'(k)}{(N-1)\ N\ (N+1)}$$

gives $$\tilde{F}(k) = F'(k) - C\left(\frac{2k-N-1}{2}\right), \qquad k = 1,N$$

where $\hat{F}(k)$ is the modified log magnitude spectrum. This technique has the general effect of reducing "noise" in the cepstrum but is analogous to the removal of G(S) . R(S) from the voice spectrum.

The pitch peak in the cepstrum is obtained by searching for a maximum in the range 3.5-15 ms (66.6-285.7 Hz). Below 3.5 ms peaks are most likely due to formant structure. Experience has shown that the cepstrum should be short time filtered with a cut off time corresponding to ½ of the pitch period since this produces the most acceptable spectral envelope for formant picking purposes. Furthermore, empirical data has suggested that the range of the first formant is approximately 250-1000 Hz so the first formant is most efficiently obtained as the largest maximum turning point in this range.

CONCLUSIONS

Typical log spectral, cepstral and smoothed spectral estimates, obtained from the in-flight recordings of Fig. 5 are given in Fig. 8. The pitch peak in the cepstrum corresponds to a constant frequency of 94 Hz. The first formant peak is clearly visible and the peak picking algorithm shows it to be increasing from 520 to 576 Hz through the time course of the phonem. In this particular example the second formant peak is well defined (1400 Hz) and the third and fourth formants may also be seen at 2250 Hz and 3100 Hz respectively. This will not always be the case though the long time spectrum of the field data in Fig. 5. does show a pronounced high frequency structure. Note that while cepstral filtering removes ripple in the log magnitude spectrum, it does not necessarily remove the peak corresponding to the pitch fundamental.

Preliminary examination of in-flight recordings has suggested that this technique is a useful one for classifying specific phonems and is effective despite the noise corruption of the raw data. The nature of cockpit communications will determine the choice of phonems which should not necessarily be extracted from the most frequently occuring words. It is expected that the emotional content of the phrase in question will play a fundamental role in the eventual choice.

REFERENCES

1. Nicholson, A.N. 1973. Aircrew workload during the approach and landing. The Aeronautical Journal, 6, 286-289.

2. Nicholson, A.N., Hill, L.E., Borland, R.G. & Ferres, H.M. 1970. Activity of the nervous system during the let-down, approach and landing: A study of short duration high workload. Aerospace Med., 41, 436-446.

3. Krzanowski, W.J. & Nicholson, A.N. 1972. An analysis of pilot assessment of workload. Aerospace Med., 43, 993-997.

4. Nicholson, A.N., Borland, R.G., Hill, L.E. & Krzanowski, W.J. 1973. Influence of workload on the neurological state of a pilot during the approach and landing. Aerospace Med., 44, 146-152.

5. Lazarus, R.S., Deese, J. & Osler, S.F. 1952. The effects of physiological stress upon performance. Psychol. Bull., 49, 293-317.

6. Williams, C.E. & Stevens, K.N. 1969. On determining the emotional state of pilots during flight: an exploratory study. Aerospace Med., 40, 1369-1372.

7. Kunoda, I., Fukiwara, O., Okamura, N. & Utsuki, N. 1976. Method for determining pilot stress through analysis of voice communications. Av. Space & Environ. Med., 47, 528-533.

8. Hecker, M.H.L., Stevens, K.N., Bismark, G von & Williams, C.E. 1968. Manifestations of task induced stress in the acoustic speech signal. J. Acoust. Soc. Am., 44, 993-1001.

9. Simonov, P.V. & Frolov, M.V. 1973. Utilisation of human voice for estimation of mans emotional stress and state of attention. Aerospace Med., 44, 256-258.

10. Flanagan, J.L. 1972. Speech analysis, synthesis and perception. Springer Verlerg., 2nd Edition, Expanded, New York.

11. Schafer, R.W. & Rabiner, L.R. 1970. System for automatic formant analysis of voiced speech. J. Acoust. Soc. Am., 47, 634-648.

12. Noll, A.M. 1967. Cepstrum pitch determination. J. Acoust. Soc. Am., 41, 293-308.

13. Oppenheim, A.V. & Schafer, R.W. 1968. Homomorphic analysis of speech. IEEE Trans Audio & Electroacoust. AU-16., 221-226.

14. Oppenheim, A.V. 1969. Speech analysis-synthesis system based on homomorphic filtering. J. Acoust. Soc. Am., 45, 458-465.

ACKNOWLEDGEMENTS

The authors are indebted to the Chairman, British Airways for the facilities to carry out this work. The study was supported in part by the Civil Aviation Authority.

Captain L.E. Hill was formerly Senior Captain of the 707 Fleet of British Airways.

DISCUSSION

VARON:
(United States)

Are you aware of the technique of using voice point in security applications, as a polygraph substitute? These voice prints have been used in security investigations for the last few years.

CANNINGS:
(United Kingdom)

The equipment in this field with which I am personally familiar does not meet its advertised specifications. However, the technique to which you refer may be of some limited applications in the current field of study.

GREGOIRE:
(United States)

Do you feel that the quality of the recording is sufficient to detect lying on the part of the pilot?

CANNINGS:
(United Kingdom)

I'm not sure about this application. We really don't anticipate lying in relation to this (landing) environment of air operations. However, with regard to quality of recording, the techniques we have described should be capable of extracting voice parameters for the purpose of classifying workload from relatively poor quality raw data.

DETERMINATION OF STRESS AND STRAIN OF AIR TRAFFIC CONTROL OFFICERS

Professor Dr-Ing. Walter Rohmert
Director of Institute of Arbeitswissenschaft
University of Technology
Petersenstr. 30
6100 Darmstadt
West-Germany

Summary

The work place of the air traffic control officer is to be regarded as a man-at-work-system. Evaluating man's task in the air traffic control system shows that air traffic controlling means a work system with "manual" performance. All three partial functions, effecting/controlling/monitoring, necessary to be fulfilled in each work system are operated by man himself.

At Frankfurt airport a lot of field studies were carried out over a period of roughly four years including research on about 115 air-traffic control-officers. Methods were developed for assessment of stress and strain. Stress is defined as all factors of work which result in reactions of the controller's receptory and effectory system. Stress leads to strain in human being not only dependent on stress but also dependent on the distinguished individual characteristics. An overlook about all methods and techniques used for assessment of stress and strain will be given.

Based on the concept of the *man-at-work-system* and the description of strain-related work contents a new *Ergonomic Job Description Questionnaire* has been developed, the results of which allow a deeper inlook into methods and techniques needed as well for evaluation purposes as for designing future air-traffic control-systems.

At least some results of the field studies in air-traffic control research are illustrated in some examples.

The workplace of the air traffic control officer regarded as a man-at-work system

The task of air-traffic control staff may be considered as a special type of service man-at-work system. The man-at-work system is defined as a model of the relations between man and his task (see fig. 1).

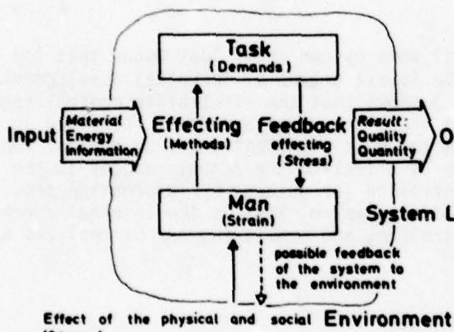


Fig. 1: Model of the working system (LAURIG, 1975)

The task of the controller is to give external instructions to a pilot in an aircraft and to make internal coordination with other controllers. Beside of input measures like presented informations about the aircraft to be guided and energy to handle miscellaneous material technical equipment, the controller needs for solving his task methods gained by education and experience. With these methods he is effecting in the work system. By solving his task demands are put on the controller which are presenting certain resistances (work load or difficulties) which must be overcome by the controller and which are meaning a feedback effecting. This feedback which influences man while doing the task, is called stress. Stress does not only depend on the heaviness or the difficulties of the task and its duration but also on the environment with its physical components (i. e. climate, noise, illumination etc.) and its social components (i. e. leadership, management relations, communication problems with other controllers or staff etc.) which are effecting as stress components as well within the working system as from outside of the system limit.

Within the man stress leads to a distinguished strain dependent not only from stress but also from different individual controllers capacities, abilities or skills. The result of the controller's task is shown in the both components of quality and quantity of the control performance. By the system limits (i. e. the controller's workplace or his functional area) the relations between man and work are marked off against the environment.

Such a general and rough description of the working system "Controller's Performance" provides for the possibility of giving the basis for detailed task analysis. With regard to the controller and his control function there are three starting points for to evaluate the controller's input and his share in the air traffic control system (see fig. 1):

- The demands of the task (and all concrete elements of the working system) which means a special kind of job evaluation,
- The qualification of the air traffic control officers, which determines differences between the methods of effecting between experienced controllers and trainees respectively,
- The capacity of the air traffic control officers (as a man-related part of the working system with regard to quality and quantity of the system's output).

To each of these three starting points to evaluate human performance there is attached the disadvantage that the feedback is neglected which influences man while doing his work. All these starting points are neglecting the evaluation of strain. If, however, an air-traffic control system contains human input and if future developments are going on to need the controller's input, too, it seems to be really necessary to consider strain put on man and its share to the system's reliability.

Evaluating man's task in the air traffic control system

The task is asking the controller for instruction on time to the pilot to provide a safe, orderly and expeditious flow of air traffic. The information for to give instructions the controller has to derive from the data about the situation in the air space as well as about the aircraft to be guided (see fig. 2): e. e. type of aircraft, call sign, route and destination desired. The function of the controller is to monitor the traffic and, if necessary, to intervene to keep traffic flow safe. The purpose is some sort of "adaptation" of air-traffic to safety criteria or avoidance of unsafe conditions. This means that information processing is the main function of air traffic control. A lot of technical equipment is needed to keep contact between the control system on the ground and the aircraft for information gathering and for giving instructions as result of processing the information. By confirming the instruction the pilot closes the communication loop between the controller and himself. The pilot changes the course of his aircraft in accordance with the given instructions. The changed situation gives new information to the controller.

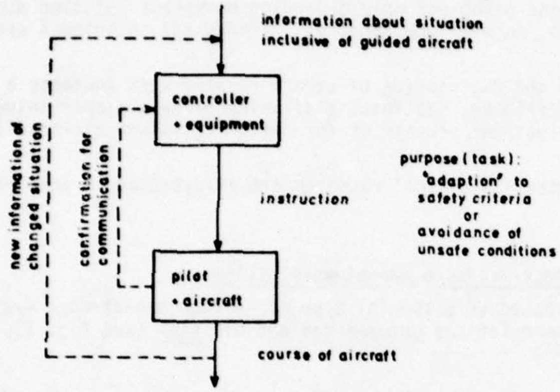


Fig. 2: Guidance of aircraft from the ground as a steering task

Information processing in the air-traffic control system is still done by man only. That means that the air traffic control task in the present state is a work system on the lowest degree of technical development as far as human functions are involved to fulfil the task. Fig. 3 shows that the air-traffic controlling means a work system with "manual" performance; all three partial functions necessary to be fulfilled in each work system are operated by man himself. In any work system the most important areas of partial functions are effecting, controlling and monitoring. By the function of effecting the actual changes in the task or in the work object will be made. The process will be controlled (or guided) by information processing. Monitoring serves to maintain the functional capacity of the system, that is the internal checking and learing of impediments. The functions of effecting, controlling and monitoring may be realized by either man or technical equipment.

WORK SYSTEM	PARTIAL FUNCTIONS OPERATED BY MAN		
'manual' performance	effecting	controlling	monitoring
mechanized performance	effecting	controlling	monitoring
automatic performance	effecting	controlling	monitoring
	PARTIAL FUNCTIONS REALIZED BY TECHNICAL EQUIPMENT		

Fig. 3
Work system with human and technical functions

Fig. 3 demonstrates three steps of technical design, logically derived from different possible kinds of work systems. The term "manual" performance means all direct human actions and also non-manual actions such as speaking. By the identification of the steps of technical design the interaction of the human operator with technical equipment is characterized revealing distinguishable types of human performance. As far as air-traffic control is concerned there are no partial functions realized by technical equipment; the controller has to perform them all three.

Technical equipment is used for presenting information (i. e. radar screen, control strip, telecommunication, telephone, instruments and scales), for pre-processing information (i. e. putting information in a more convenient form for handling by man). Without the radar screen and some indicating instruments and scales all technical equipment designed for picking up the information is also destinated and used to hand over the information. The guidance of aircraft from the ground is a control task in which feedback only serves to control the execution of the instructions and give new information about the changed situation for future actions of men. This can be called as some sort of a"manual" steering performance.

Information processing is devided into two steps (see fig. 4). In the first step the incoming information is analysed and evaluated. For this purpose the controller needs storage for the evaluation of situations which comes from training and experience. Then the controller has to select a proper answer or action out of the store. But the controller is not only concerned with processing information about the situation in

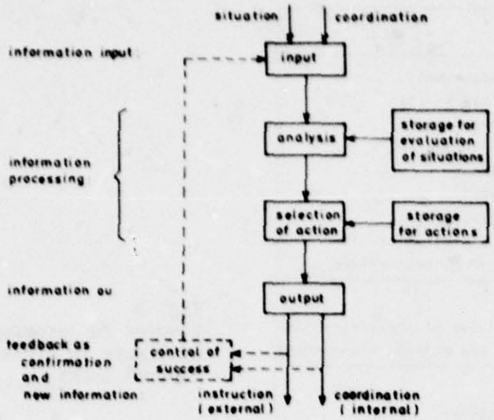


Fig. 4: Information processing by the air-traffic control system

the air, leading to instructions to the aircraft. He must also process internal information which comes from coordination between different controllers. This additional task arises from the necessity of devid-ing the whole task of the air-traffic control system into separate jobs which can each be done by one man only. Coordination is an internal task in air-traffic control resulting from the internal organization. One cannot neglect it because it occupies much of the capacity of the controller and his assistant.

Methods for determining stress and strain in air-traffic control tasks

As explained in fig. 1 we have to distinguish between "stress" and "strain" while analysing man at work relations. To render precisely the nearly synonymous meaning of these two terms stress is defined as all factors of work which result in reactions of the controllers receptory and effectory system. Thus stress leads to strain in human being. However, the amount of strain is not only a function of stress but depends on the capacity of the individual, i. e. a certain amount of stress will lead to various amounts of strain in different human beings corresponding to their individual capacities (see fig. 5).

Fig. 5: Relationship between stress and strain

$$\text{Stress} = f\,(\text{partial stresses}) = g\,(\text{stress-intensities, stress-duration})$$

$$\text{Strain} = h\,(\text{stress, individual characteristics})$$

Fig. 6: *The concept of stress and strain*

For some different types of human work the functional relationship between stress and strain which is shown in fig. 6, has been described on the basis of the results of research work. We picked up this idea and technique and did a lot of field studies over a period of four years with German air-traffic control officers at the airport Frankfurt. Fig. 7 and 8 are presenting an overlook about the methods and techniques we used.

STUDY	OBJECT OF INVESTIGATION	METHODS
Workplace	Dimensions of workplace Devices of work and information used Body position demanded Environmental conditions (noise, illumination, climate)	Description Drawing Photographic documentation Physical measurement
Process	Intensity and Chronological, spatial and information-technical flow of defined work-content (also task elements, governing time) Smoking Coffee drinking	Description Graphic representation Use of standard analysis charts File analysis Tape record commentary chronological covering of telecommunication frequencies Stop watch studies Stop calculator studies Work sampling Automatic registration Continuous film records Sampling film records Telemetric coding scaling task difficulty coding control strips coding controlled aircraft Registering frequency of coordination and manipulating the control strips Uninterrupted working time at position Cumulative time Shift analysis by self-rating
Motion analysis	Sequence of hand (and body) movements of the controller	Use of standard analysis charts Use of predetermined motion time systems File analysis
Information-content analysis	Density of information flow (syntactical content of information) Pragmatic content of information (aircraft- and/or controller-related operators and logical operations Traffic and its density Strategies in control task	Study of flight control strips Photographic radar recordings (each 90 sec. time of exposure) Radio-telecommunication transcriptions Algorithmic transcription LJAPUNOW-language Flow diagrams Graph representation (GERT) Interview Number of conflicts
Time-budget study	Socio-demographic variables Daily distribution of work, rest, official rest-pauses, time for going to work an home, sleeping, spare time	Interview Daily time analysis by self-rating

Fig. 7: Studies for assessment of stress in air-traffic control tasks

STUDY	OBJECT OF INVESTIGATION	METHODS
Physiological studies	Intensity and flow of strain	Heart-rate measurement Arrhythmia of heart-rate Long-electrocardiography Respiration measurement Tremor measurement Electromyography Electroencephalography Catecholamine secretion in controllers
Applied Psychological studies	Job satisfaction	Interview Study of attitudes of the controllers towards work and working environment
Studies of work medicine	Disturbances of health related to the controllers' profession	Interview Time-budget studies Long-electrocardiography Experimental shift work

Fig. 8
Studies for assessment of strain in air-traffic control tasks

Fig. 9 shows the complete experimental set-up used in the field studies at Frankfurt airport. It comes out from the graph that the methods and techniques used in these field-studies were partly classical ones, partly they had to be developed for the purpose of the investigation. The upper part of the graph shows the judging controller who is operating a four-item telemetric coding device with different scales for assessment of the difficulty of controller's work, the number of incoming control strips and the number of aircrafts under control. He is also giving additional commentaries about the general or specific work situation to the tape recorder. All other studies for assessment of stress mentioned in fig. 7 were done by one of our staff who is mentioned in fig. 9 with the task of supplementary work study. Some of the methods used were working automatically (i. e. picking up radio-telecommunication on magnetic tape, radar pictures). All time-budget studies were done within the periods of brakes or between two shifts. Most of the stress variables were transmitted telemetrically; the lower graph in fig. 9 shows the total telemetric receiver set-up for the incoming stress and strain data and the data storing set-up.

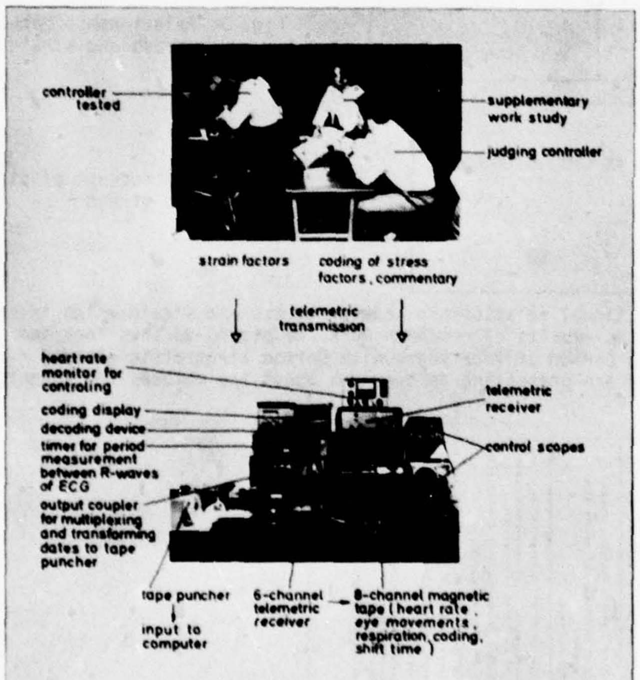


Fig. 9: Complete experimental set-up

To transmit the physiological variables for assessing strain be used a multi-channel telemetric system. In fig. 10 the arrangement of electrodes for deducing the different electrophysiological variables and the transmitter is demonstrated. Also the attached respiaration sensor can be seen. While some of the strain variables were transmitted continuously during the working time on the position (heart rate, respiration rate, electromyogramme, electrooculogramme) other variables were picked up for comparison reasons two times: first, just before the officer is starting his work at the position, second, just after having finished the work at the position (tremor measurement, catecholamine secretion, electroencephalogramme). All other studies mentioned in fig. 8 were performed during brakes or between two shifts.

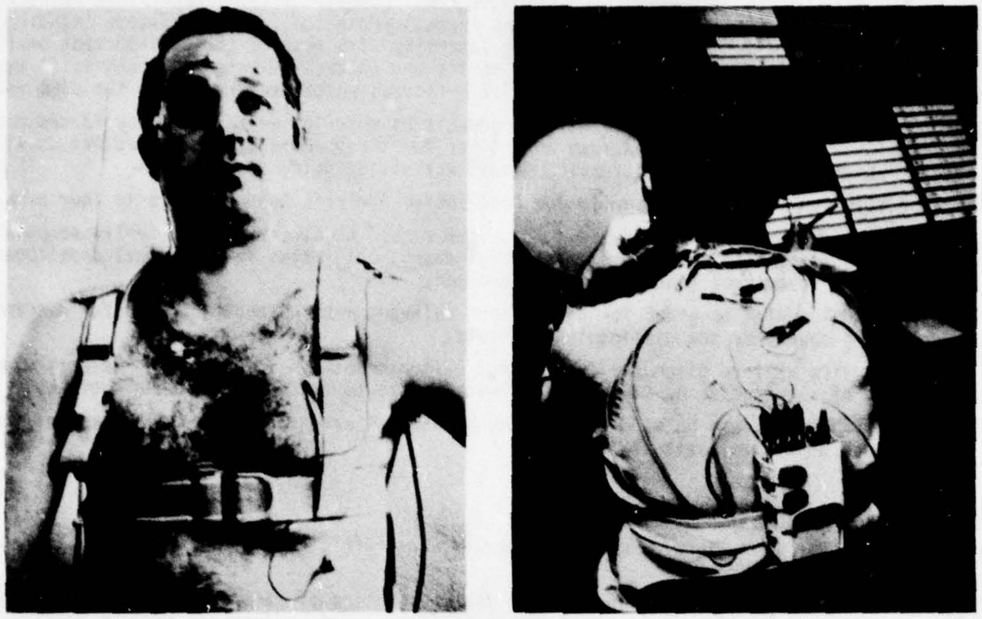


Fig. 10: Electrodes placement and telemetric transmitter

Determination of stress of air-traffic control officers by position analysis

Based on the concept of the man-at-work system (see fig. 1) and the description of strain-related work-contents a new Ergonomic Job Description Questionnaire (EJDQ) has been developed (ROHMERT, LUCZAK, LANDAU, 1975). This method allows a very detailed analytical job evaluation. A trained job-analyst analyses the activities of the operators in air-traffic control systems with respect to very detailed job elements (the job evaluation procedure contains 390 items).

Various rating scales, such as importance to the job, frequency, probability, amount of time (spent on the job element), are used for the different job elements.

The demand structure of air-traffic control tasks can be evaluated by grouping together related items with respect to stress. The rating scale values of item groups are summed up and plotted columnwise. Item groups are derived straight forward from the demand-orientated chapters of the Ergonomic Job Description Questionnaire. However, they can be derived as well from the results of a factor analysis of the Ergonomic Job Description Questionnaire data.

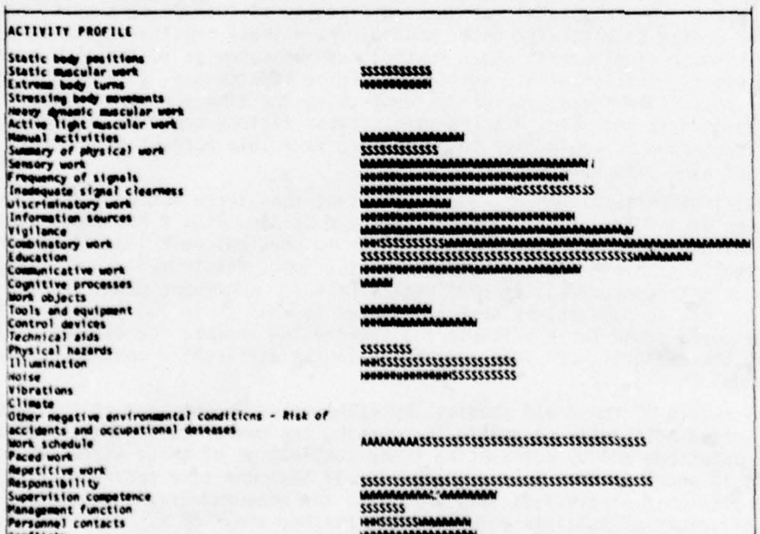


Fig. 11
Activity profile of an air-traffic control officer in the pick-up position at Frankfurt airport

Fig. 11 shows the activity profile of an air-traffic control officer in the pick-up position of Frankfurt airport approach control. Characteristic demands are combinatory stress, education and vigilance. Besides that, stress caused by work schedule and responsibility are important stress factors.

The profile shows high work load especially for sensory systems, caused by bad-conditioned and -adapted work means (i. e. display and information output). Stress factors of information input are given by the frequency and clearness of signals and information sources. Whereas frequency of signals cannot be changed significantly, clearness of signals - the design of control systems in general - might be improved.

The analysis of physical working conditions shows good thermal conditions (air temperature about 25°C, relative humidity 46-49 %, wind velocity 0,2 m/sec, no thermal radiation), but inadequate lighting of the darkened landscape office is given and high noise intensity with respect to the important part of mental load. A noise level of 53 dB (A) during night shifts and 64 dB(A) during other shifts is caused by loud-speaker,transmitted radio-telecommunication and telewriters, which are running in the same room.

Characteristic physical stress factors are static components of muscular work, caused by forced body positions and continuous attention, and extreme body turns for the purpose of work coordination with other operators. Furthermore, there are stressing manual activities using tools.

The importance of this very detailed Ergonomic Job Description Analysis Approach lies in four aspects:

- First, main stress factors are indicated; with regard to stress one may derive proposals for design activities in air-traffic control tasks; priorities for technical development in air-traffic control functions may be deduced.
- Second, advice will be given for to find out relevant methods and technique for determining strain caused by the distinguished stress.
- Third, advice will be given for to find out suitable methods for personnel selection and training of air-traffic control officers with regard to the specific demands of their job.
- Forth, comparisons can be made between the demands of air-traffic control tasks and the demands of other professional tasks.

Some results of field studies in air-traffic control

Within a period of roughly four years we studied some 115 air-traffic control officers at different work places at Frankfurt airport (ROHMERT).

We found irregular distributions of the duration and the frequencies of task elements. Significant differences in these findings are coming out if the tasks during rush hours and periods of less traffic are compared which means that the controller is using more standardized strategies under the pressure of rush hours. The amount of information processing per time is twice higher in the approach control compared with the area control, although the average number of aircrafts under control could not be distinguished. This means that with regard to the same traffic, the control task in the approach control is more difficult than in the area control. This result was confirmed by heart-rate findings: Heart-rate was higher in the average in the approach compared with area control.

Within our field studies we concentrated our activities to two important strategies: First of all, we developed a multi-dimensional measuring concept, which means that we measured several variables of stress as well as of strain (see fig. 7 and 8). Furthermore, we measured most of the stress and strain variables as time series, which means that we intended to correlate a time series of a distinguished stress factor with the time series of another stress factor or with time series of dinstinguished strain factors. By doing so one gets a deeper inlook to the relationship between stress and strain. When testing correlations between time series, the effect of serial correlation has to be considered. The calculation of regression coefficients for the description of the relationship between serial correlated variables requires assumptions about the validity of the least square procedure. After taking into consideration further analysis of the residuals and serial correlation effects a combination of stress factors can be used to determine a distinguished strain factor or a combination of strain factors respectively.

We found significant relationships between the results of rating of difficulty of controller's work, coding of incoming flight control strips and coding of aircrafts under control. From these results it can be summarized that the stress variable "number of aircrafts under control" can be taken as a stress measure which describes the time-related stress fluctuation with a sufficient degree of accuracy. This is the reason that we took this very simple measurable stress factor for describing the time series of stress in air-traffic control. This simplification does not mean that the other stress factors could be neglected. As far as further field research is concerned we would like to propose to take into account also other stress factors beside of the number of aircrafts under control.

Strain can be evaluated only in a multi-dirensional manner, due to the fact that there could be different bottle-necks caused by different human organs or abilities, capacities and skills. Fig. 8 had shown some examples for bottle-neck-orientated strain measures, related for example to physical work load (i. e. heart-rate), mental work load (i. e. arrhythmia of heart-rate, electroencephalogramme, electromyogramme), sensory work load (i. e. electrooculogramme, electromyogramme), or emotional stress (i. e. tremor measurements, catecholamine secretion) respectively. But it came out by several studies as well as in the laboratory as in the field that already heart-rate could stand for a suitable and integrating measure for evaluation of strain if one succeeds in describing the differnt work load components in the air-traffic control tasks by suitable stress measures. This is shown in fig. 12.

Fig. 12 is ullustrating an important result of the field studies. By mathematical connecting of the three measured time series of stress (the rated difficulty of controller's task, the number of aircrafts under control and the number of aircrafts expected) and by correlating these combination of three stressors with the strain measure of heart-rate, it is shown in the upper curve of fig.12 that the time series of strain can be predicted. The course of the predicted strain fits very well with the measured strain which is marked by the small circles. The coefficient of multiple determination reached about 80 %.

Fig. 12 indicates, too, that heart-rate is an integrating measure of strain. Heart-rate is not only influenced by the mental load of the controller, which may be evaluated by determining the number of aircrafts under control or the rated difficulty in fulfilling the task, heart-rate is also influenced by the more emotional stress which can be evaluated by the number of aircrafts which will be expected. The number of expected aircrafts is reaching its maximum before the number of aircrafts under control or the rated difficulty is showing maximal values. Due to these emotional reactions an increase in the intentional basic tension of the air-traffic controller can be expected. Due to this hypothesis heart-rate must show an increase also with the increasing number of aircrafts expected. Fig. 12 shows the highest values of heart-rate not only at the peak of mental load but also at the peak of emotional load.

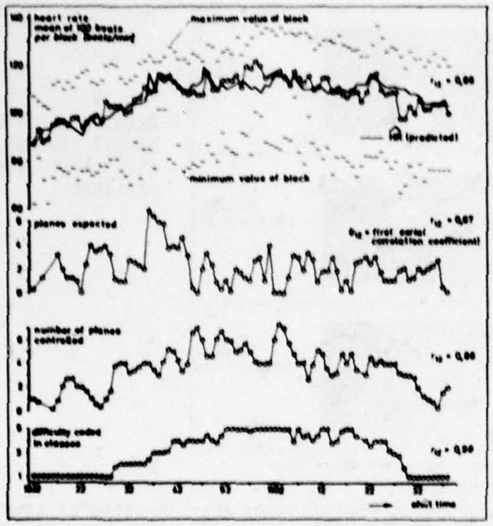


Fig. 12
Variation of heart-rate and stress factors (approach control, pick-up position)

By some further additional experiments it could be shown that a good prediction of the strain measure heart-rate could be gained also by evaluating the stress measure of number of aircrafts under control onlys. And as far as heart-rate is influenced moreover not only by mental but also by emotional work load, we preferred this measure for evaluating strain. This does not exclude that other strain measures allow a suitable description of specific components of the work load of air-traffic controllers. This might be shown in some examples of measurements of tremor activity as well as catecholamine secretion.

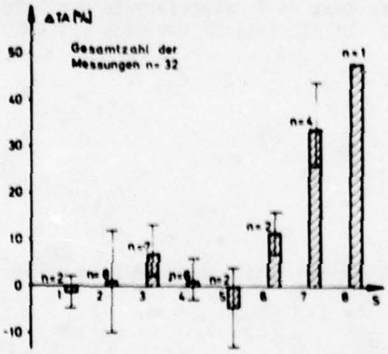


Fig. 13
Increase of the amplitude of tremor movements (ΔTA) of the hand dependent from the task difficulty (S)

Fig. 13 shwos that with increasing difficulty in fulfilling the controller's task the amplitude of tremor movements, that means the micro-vibration of the lower arm and hand system, will be amplified significantly. Similar results we only found in research on the situations of students examinations and doctor-thesis situations. Also the increase in catecholamine secretion is as high as in emotional situations of exceptions, like the examination for getting the driving licence for motorcars. These values are five times as high as the values of the teachers who are leading the examinations.

The results of measurements of heart-rate, tremor activity and catecholamine secretion show that the air-traffic control task is very high emotional stressing. The effect of this stress component can also be compared with exceptional situations in the daily or professional life. But contrary to these which will occur only a few times in man's life, the emotional stressing situations of controllers will be repeated day by day.

All the results gained in the experiments with the total of 115 air-traffic controllers from Frankfurt airport, were evaluated with the purpose of getting a model for predicting the changes in strain dependent from *the number of aircrafts under control* and *the cumulative time* of the controller's work. The model shows that the intensity of work and the duration of work are influencing the increase in strain in an exponential manner. But this is valid only if the number of aircrafts under control is exceeding a distinguished limit of tolerability. If the intensity of work is lower than this limit the duration of work shows no significant influence to strain.

We used the results of the model for predicting a suitable work load in air-traffic control tasks. Also proposals were derived for optimal schedules of working time and rest pauses. With respect to strain we made further some proposals for schedules of the shift organisation.

Some results of the time-budget studies

Increasing strain is not only influenced by the intensity of work but also by its duration. Therefore, we studied also the influences of the duration of the shift and of the shift's position within the 24 hours of the day. By these time-budget studies the controllers rated for each half hour of the day whether they had a working time, a sleeping time, a leisure time which could be spent in full social freedom or which was related to other daily activities. Last not least the controllers had to state whether they were on the way to their work place or back on the way home. Fig. 14 shows the results in relation to different types of shifts and for off-duty days.

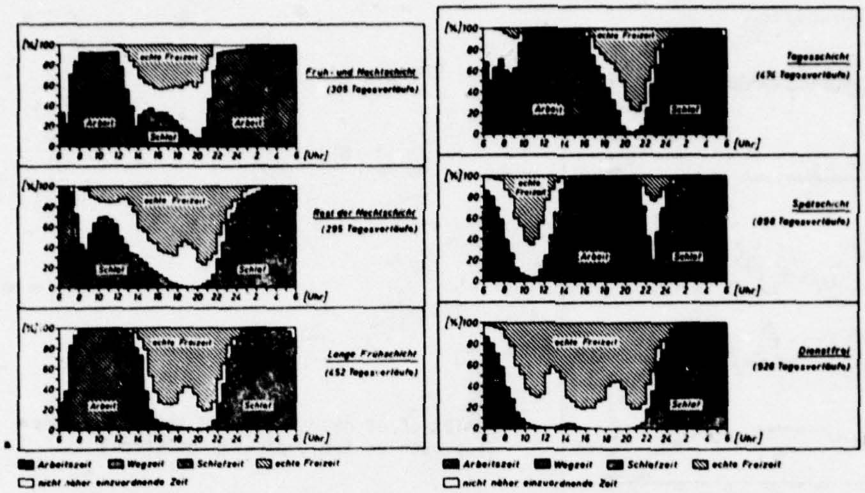


Fig 14
Results of time-budget studies of air-traffic control officers

There are some remarkable results. First of all, it is outstanding that the time of starting with sleep is very constant and nearly not related to type of day shifts. This means that the length of sleep is only related to the time of waking up. In the combinated early- and night shift there are two sleeping periods which are characterized not only by less but also by worse sleep. Alos the leisure time which is at the free disposal, can be evaluated in relationship to its location within the 24 hours period of the day. It can be seen furthermore that the time for travelling to the work place and back home is not independent from the type of the shift.

The results of the time-budget studies may lead to the conclusion that it is necessary to evaluate also time-budgets, if man's reliability is one of the bottle-necks of the elements of the work system.

References

KIRCHNER, J.-H., LAURIG, W.
The Human OPerator in Air Traffic Control Systems.
Ergonomics 14 (1971) 5, 549-556

KLIMMER, F., AULMANN, H. M., RUTENFRANZ, J.
Katecholaminausscheidung im Urin bei emotional und mental belastenden Tätigkeiten im Flugverkehrs-kontrolldienst.
Int. Arch. Arbeitsmed. 30 (1972), 65-80

KNAUTH, P., RUTENFRANZ, J.
Untersuchungen über die Beziehungen zwischen Schichtform und Tagesaufteilung.
Int. Arch. Arbeitsmed. 30 (1972), 173-191

ROHMERT, W.
Ermittlung von Belastung und Beanspruchung der Fluglotsen in der Flugverkehrskontrolle.
Industrial Engineering 2 (1972) 1, 23-30

ROHMERT, W.
Ergonomische Beurteilung der Belastung und Beanspruchung von Fluglotsen in der Flugverkehrskontrolle.
Industrial Engineering 3 (1973) 2, 99-110

ROHMERT, W.
Psycho-physische Belastung und Beanspruchung von Fluglotsen.
Beuth-Vertrieb, Berlin/Köln/Frankfurt 1973

INSTRUMENTS AND METHODOLOGY FOR THE ASSESSMENT OF PHYSIOLOGICAL COST OF PERFORMANCE IN STRESSFUL CONTINUOUS OPERATIONS - THE AIR TRAFFIC SERVICES TOWER ENVIRONMENT

Jacques Soutendam, Major
Defence and Civil Institute of Environmental Medicine
1133 Sheppard Avenue West
P.O. Box 2000, Downsview, Ontario
M3M 3B9
Canada

SUMMARY

A total of 39 females and males were used as subjects in a study designed to test the general validity and utility of methods and instruments of potential use in the determination of the physiological cost of work performance in stressful environments.

A second goal of the study was to attempt a preliminary test of the hypothesis that a form of employment (Air Traffic Services) traditionally considered stressful, was significantly different from general forms of employment traditionally considered less stressful. To test this hypothesis 23 people employed as air traffic control tower personnel were slated as the Experimental group and 16 subjects employed in an operationally oriented research facility served as the Control group.

Within the parameters of this research design, consistent statistical significance was established when the entire sample was re-categorized on the basis of subjective sleep adequacy assessment rather than by job description. Further research is indicated with well defined physiological categorization measures to obtain more definitive answers concerning job stress and physiological cost and the methodology development necessary to support general research efforts in this field.

INTRODUCTION

Within the past decade much information has become available concerning problems arising in connection with the performance of the human sub-system in the extremely complex Air Traffic Services Man/Machine Systems environment. Some of the literature is strongly emotional in character, some of it is observational in nature, and the majority is scientific or technical in content. This wide range of information and conjecture has not always provided clear solutions to perceived or demonstrated problems and it is evident that a restructuring of the systems philosophy for research is required. In addition to the difficulties caused by the complexity of the ATS system itself, the major research problem is the human in that system. With the human being in the system is associated a genetic predisposition; the behavioural effect of a lifetime of myth and morality and an exquisite individuality that make "the person" a challenge to quantifiable description.

There is evidence that unrealistic output demands are being made on the ATS human sub-system and that, in addition, it is believed that an increase of controller responsibility will no longer mean a commensurate increase in Air Safety. Knowledgeable people feel that we have arrived at a point in ATS function where more information and responsibility must be relegated to the pilot.

In contrast to the relatively rigidly structured and powerfully standardized environment in which the pilot functions as the human sub-system, the ATS control tower environment displays no strong measure of uniformity in its design, which suggests that much of the information gained in any environmental study research project may be site specific rather than solely job specific. Within such restrictions, this paper and other observation and interview findings must be read.

It has been suggested by some that ATS needs a miracle badly if some degree of sanity and balance is to be brought into an environment that is almost uniformly considered highly uneconomical in terms of physiological and psychosocial performance cost. The fervent hope for a miracle springs from the fact that much of the scientific and pseudo scientific research carried out in the ATS environment has failed to yield findings that unequivocally demonstrate significant differences between ATS controllers as a sample group and any selected non-ATS control group.

This must not be interpreted as necessarily suggesting that ATS is a non stressful environment, but it does mean that within the specific research program parameters, no significant differences could be demonstrated, and in the light of these non-findings, a realistic perspective must be kept when focussing on the need for further research. Whether or not such future research programs do in fact occur, there are groups that are planning action that they expect (hope) will allow them to come to grips with the ATS stress problem.

Many of the discussions in current literature concerning the ATS problem are traditional in tone and refer to periods of exposure to intense job related stress and associated safety implications. Illness and absenteeism rates for controllers are cited and often the need for increased job satisfaction is stated as a potential cure-all. Despite this, there is compelling evidence suggesting that job satis-

faction is in fact not a concern in ATS. An extensive study was conducted by the Aviation Psychology Laboratory of the Civil Aeromedical Institute of the Federal Aviation Administration to determine job satisfaction among a large number of controllers. Conclusions of the U.S. study, drawn from the information received indicate that on a scale of 0 to 100 percent, controllers average a 91 percent level of job satisfaction. An abbreviated Canadian study, admittedly involving a much smaller sample consisting of tower controllers only, indicated a nearly equivalent 88 percent job satisfaction level. An additional important fact, emerging from both studies, is that controllers like, and in fact prefer, high levels of job activity. These findings raise some questions concerning the accepted belief that the stress of high work loads are necessarily detrimental to the self perceived well-being of the controller.

During a recent meeting hosted by the American Academy of Stress Disorders, a plan was tabled on the suggestion of Dr. E. Eliot Benezra, a psychiatrist who has worked with a number of air traffic controllers. Dr. Benezra feels that a Crisis Intervention Team (CIT) should be available at all times to practicing controllers. The Team might consist in part of controllers with an aptitude and training to understand the emotional aspects of their colleagues.

In the light of scientific study findings and the crucial career impact of the stigma of psychiatric treatment there is some question about the validity and the wisdom of such a dramatic, and to the controller possibly traumatic, approach.

Another approach forseen by some as the solution of the human problem in the ATS man/machine system complex is the radical increase in instrument technology. It is suggested that an advance in the state of the art of raw information presentation systems will improve information processing ability by the human sub-system and thus enhance aviation safety.

In terms of reducing ATS job stress this could prove to be a questionable course as well. From what is known today, there are strong indications that the limits of man's mental capacity to absorb and process information are being reached, and that people capable of consistently handling present levels of complexity are becoming more and more rare.

In ATS, the introduction of more powerful and complex information presentation systems will require the controller to assimilate even larger quantities of data in a severely limited and critical time frame and it is possible that his or her physical and intellectual limits are being reached and will soon be traumatically exceeded. The other associated fact that ATS problem solvers must come to grips with is that a second complex man/machine system is involved in this already saturated system; that is the pilot/aircraft unit. It is suggested that a possible way of decreasing the mental loading on the controller would be to make cockpit instrumented aircraft proximity information available to the pilot as well as to the ATS controller during the critical final phases of flight. It is felt by some critics of the North American ATS Systems that this is not only desirable but mandatory in the interest of flying safety and reduction of human sub-system stress. The first step in future projects must be to clarify the questions and formulate legitimate empirical means to test the hypothesis. We need not re-make all of the mistakes.

STUDY INTENT

Concern in this study lies with the traditional methods of laboratory research, the methods needed in the conduct of field research and the methodological issues and strategies appropriate for linking the domains of scientific activity in the field and in the laboratory, specifically in relation to predicting of physiological cost of human performance during continuous work under stressful conditions.

Although the questions asked by the researcher or user serve as the input to the research flow model, and are the driving force in all research, and provide the hypothesis to be tested in more formal scientific terms, they are never more critical and practical then when they provide information of a predictive or potential control nature for a consumer community. Initially the problems must be defined in such a way as to be meaningful to the user and researcher alike. Once this has been done, there may be information available by which the problem can be solved, either uniquely or through low cost tests of alternate solutions. Should a major research effort be required to generate the information needed, several strategies may be employed.

Any strategy or technique used, must meet specific researcher and user criteria in that it must be:

a. Meaningful (to the consumer)

b. Standardized (results evaluated against defined standards)

c. Reproducible (results should be consistent when repeated)

d. Interpretable

e. Efficient (economic expenditure of investigation effort)

To offer research findings leading to applications for a meaningful solution to the user, the research design must be sufficiently flexible in its applicability to allow researchable universe substitution. In selecting Air Traffic Services (ATS) for scientific investigation, recognition was taken of the amazing strides made by civil aviation and how all forms of this transportation mode have affected the course of history by their strong influence on commercial, cultural and diplomatic relations. Efficiency effectiveness and safety of air transport is in large measure due to the high professional qualities of the pilots and ground staff, such as Air Traffic Controllers. Among the ground services, ATS is the only one to have direct effect on the conduct of flight. The ATS man/machine environment systems complex has perhaps as its most valuable link the human being, for man is still the final and key element in decision making.

In the event of an emergency, with the pilot having the responsibility for the aircraft in difficulty, only the air traffic controller can ensure that it has unobstructed air space and a cleared flight path.

The air traffic controller must continue to control the movement of many aircraft simultaneously, often as many as 15, whether or not an emergency is in progress. Each decision must be taken without delay and often without the possibility of recall.

Studies (1, 2, 3) indicate a high occurrence of cardiovascular and gastrointestinal disorder and suggest that top physical and mental condition are vital requirements for the air traffic controller. Staff are required to undergo stringent medical examinations identical with those given for professional pilots and these are repeated at relatively frequent intervals. A phrase often used by ATS staff to describe their work is "unavoidable high nervous job tension", which it is felt can be compensated for by harmonious working environments, suitable working hours and periods of rest (4).

It is suggested that it is important that an above average to high standard of living should be provided for ATS personnel to relieve them of as many of the day to day worries as possible. This is seen as necessary because the air traffic controller's slightest error could cause considerable loss of human life and property. His professional occupation and living standard is at stake with each decision he makes.

What emerges strongly from all conversations and ATS seminars is the desire and felt need for special status. Since its founding in 1961, the International Federation of Air Traffic Controllers Association (IFATCA) has sought to protect and promote the profession of air traffic control. Through IFATCA access can be gained to the International Civil Aviation Organization (ICAO), the International Labour Organization (ILO) and many other specialized organizations such as the International Federation of Airline Pilots' Association (IFALPA). It is generally and genuinely felt by ATS that only recognition of the realities of the pressures of this occupation and rights of the profession will ensure that civil aviation can continue its development in the most complete security, safety and user confidence possible.

Within the framework of this research project and on the basis of the wealth of literature available on stress related behaviours and performance in ATS, it is suggested that the ATS proposals for handling the problems of stress may have inadequate scientific backing and as such are not yet adequately supportable.

1. A meaningful definition of stress has not yet been made, although Selye (5) offers compelling evidence of it being a non-specific biochemical response.

2. Biochemical stress indices historically used in research, although suggestive, do in fact not reflect the adverse nature of the stressor especially over long and/or repeated exposure.

3. Although special status and financial reward may make working under stress more profitable and financially acceptable, it will not alter the stressful nature of ATS employment which exacts its toll on the physical and mental being of the Air Traffic Controller.

4. The question of ATS physiological job cost has not been investigated in a manner adequately meaningful to the user.

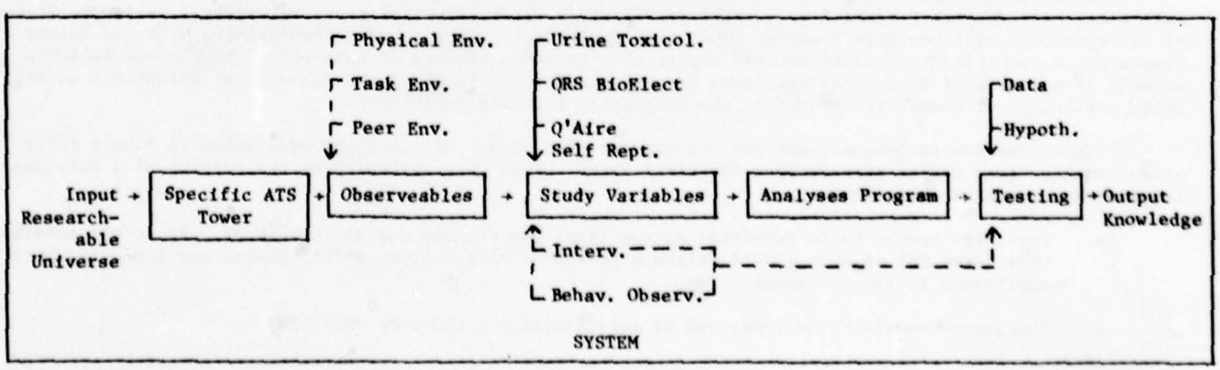
What would be of significant benefit to ATS, is research directed towards finding the possible cause factors that have established ATS as a high risk form of employment (6, 7, 8, 9, 10) and programs geared to experimental intervention aimed at either reducing the stress without compromising necessary work output and quality or at teaching the human being to cope better with the stress.

The comparison group in this project was selected from among federal government employees who are staff members of the Defence and Civil Institute of Environmental Medicine (DCIEM), a facility engaged in scientific research and operationally oriented application of findings. It is a fair assumption that performance demands are as high and failure penalties as personally critical as in the ATS environment, although the risk of psychiatric and clinical breakdown to the same degree as in ATS has not been demonstrated for DCIEM staff. On the contrary, the risk appears to follow the predictable general population pattern (11).

The descriptive model in this study cannot include all operational elements because of the complexity of interrelationship between the work environment, the task itself and the man performing the task. Study objectives are preliminary as are the analysis, the identification of initially critical variables, and the testing of hypothesis dealing with the relationship between sub-strategies, in terms of physiological cost. The systems approach is thus incremental in concept and is a first attempt to formulate and operate within the step-wise flow model as indicated in Figure 1.

FIGURE 1

RESEARCH PARADIGM:
ATS/CONTROL CONTINUOUS OPERATIONS PHYSIOLOGICAL COST ASSESSMENT SCHEME



RESEARCH METHOD

The general objective of this basic research program is to develop and evaluate a brief series of research methods and strategies for future application to the determination of the physiological cost of employment, which is both demanding and continuous in nature, while not impinging significantly on the ongoing workload of the subject group. Specific objectives are the determination of the internal and concurrent validity of measures and instruments used and the evaluation of research sub-strategies varying in style and content.

The occupational arena selected for this preliminary study was the airport control tower, which is a complex man/machine systems environment, considered stressful as a result of the substantial psychomotor performance demands made continuously on the human operator in the system. The sample of people chosen included both female and male controllers, capable of working all tower positions, and with a precisely defined task. The ATS tower facilities selected were Toronto International (CYYZ) and Buttonville (CYKZ) airports. The controls consisted of female and male personnel employed at the Toronto based Defence and Civil Institute of Environmental Medicine (DCIEM) an operationally oriented research and educational facility.

The control group in this study was selected on the basis of having a similar history to that of the sample. The significant difference being the assessment of control group employment as traditionally not stressful. This satisfies the system concept requirement that the output at any particular time, in any particular run is dependent only on sample condition and not on its future to that time.

All measurements and observations were carried out during a daytime working day, in the interest of sampling consistency.

All subjects were volunteers and because of the very small population to draw from, no limits on age, sex, marital status or length of employment were applied. With two exceptions the entire group fell between the ages 23 to 38 years of age, 80 percent were married, the group was predominantly male, and all but 5 of the group had been active in their present work between 1 to 10 years.

Data collection, reduction and analysis occurred over a seven month period and involved the following:

a. The development and assessment of a methodology for the screening of urine for toxicological substances present.

b. The application, data collection and analysis of a series of self reporting questionnaires allowing subjective psychophysiological state assessment.

c. The development and assessment of a methodology for continuous electrocardiogram monitoring for sinus arrhythmia analysis as a mental activity index.

Additional observations were made in an attempt to gain personal experience and knowledge which could prove to be of supportive value and these included the following:

1. in-depth interviewing following the duty period, and

2. continuous unobtrusive behaviour observations during duty.

In each of the methodological strategies, different measuring devices and instruments were used and to enhance the reader clarity of this study as well as to facilitate potential user orientation, methods and instrumentation information is separated and is detailed appropriately.

The private interviews were carried out along structured lines, but were sufficiently flexible to allow the subject to digress in a guided manner. The intent was to obtain personal information allowing the author to make an informal subjective assessment of:

a. the working environment quality, and

b. psychosocial environmental needs.

TOXICOLOGICAL ANALYSIS

CHEMICAL COUNTERMEASURES

The toxicological analysis in this study attempts to establish whether or not an individual participating may have subjected himself or herself to voluntary pharmacological countermeasures. In future studies it is envisioned that positive findings could be correlated to other sub-strategy results with the intent of presenting a specific objectively defined aspect of a reliable history of behaviour. Simple and reliable methods of analysis of biological specimens are not readily available to the non-medical scientists so the DCIEM toxicological chemistry laboratory was approached for this assistance.

Since invasive techniques were not planned for this study, all subjects were asked to submit three urine samples, each sample representing the first urine voided upon awakening on the morning of a duty day. Samples were required as follows:

a. the first sample to be submitted on the first day of duty during this study. If at all possible, this first day of duty should follow a period of rest between shift changes particularly where applicable to the ATS group.

b. the second sample to be submitted at approximately a mid-duty shift point.

c. a third sample to be submitted at any time within a two week period following collection of the first sample, as long as the sample was taken while on duty.

On submission, the urine was immediately frozen and transported to the DCIEM laboratories where it was analyzed according to established methods.

Research literature reveals a perplexing array of screening methods. A requirement of this study was to produce reliable, meaningful data, but hopefully it was to produce a simple and relatively speedy procedure as well. Following consultations with DCIEM toxicology specialists it was decided to use this study to evaluate the feasibility of a multi sample modeling, satisfying the listed study research design parameters.

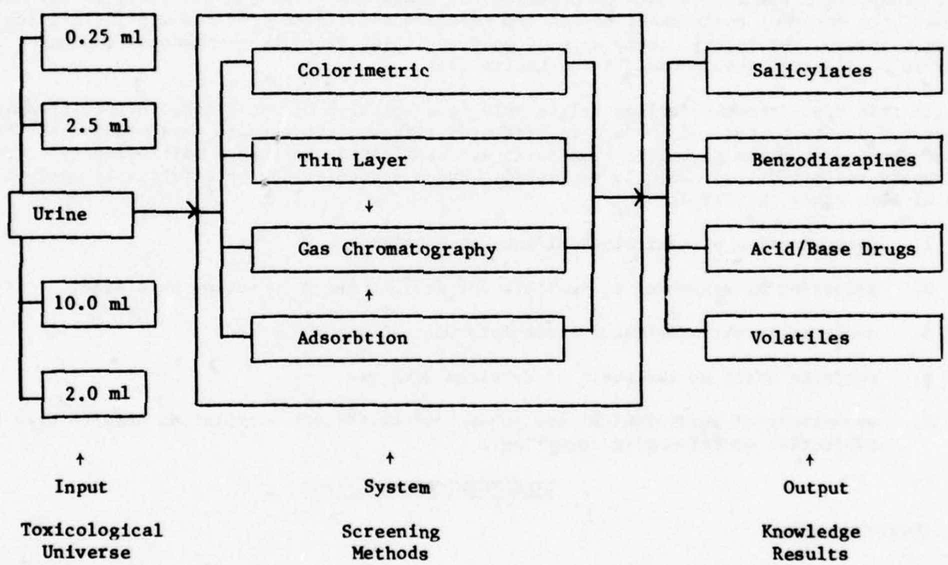
It was apparent that the types of methods planned were elegant and precise and not necessarily within the limits of routine experience, and the reader is reminded that this study is not concerned with industrial toxicology and no analytical efforts in that discipline are anticipated. Detailed descriptions on methodological techniques are not offered as it is considered within laboratory expertise and not part of this study. However a brief explanation of each technique is given.

For sensitivity drug screening, any isolated procedure is considered inadequate for identifying a compound, thus a combination of procedures, and subsequent procedure coupling is applied. It is considered that only in this way can sufficient information about a substance be provided, guaranteeing the reliability of a statement made about it. The test combinations are schematically presented in Figure 2.

Urine screening in the study subjects was performed, by sample in the following manner:

FIGURE 2

THE TOXICOLOGY SYSTEMS MODEL



1. Salicylic acid screening required a 0.25 ml urine sample. The method used was that of Irving Sunshine, (12) Salicylate type A procedure. Results are reported accurate within ± 5 percent, determine urine salicylate concentration, and thus identify the use of non-narcotic analgesics in the salicylates group, such as the common aspirin. Precise quantitative analyses were not considered essential for this study.

2. Screening for the benzodiazapines class of drugs required a 2.5 ml urine sample, analyzed by the method used by Dr. B.N. Kapur, involving thin-layer chromatography. Ready-coated plates with thin layer silica gel as an adsorbent are submerged in moisture chambers containing a solvent which migrates through the layer. Spraying with reagents visualizes the drug on a chromatogram. The use of tranquilizers such as Librium and Valium are indicated. Pending general availability of the reference publication, Appendix "A" to the study briefly details the Kapur method.

3. Sedatives and hypnotic classes of drugs required a 10.0 ml urine sample which was analyzed for the acid and basic drugs. The method used was that of John M. Meola and Michael Vanko, (13, 14) where drugs are absorbed from urine onto charcoal selectively eluted, and identified by thin layer chromatography. Positives are further investigated by gas liquid chromatography. The system used is the 7620A Research Chromatograph, 12 module complex, manufactured by Hewlett-Packard.

4. Determination of volatiles required a 2.0 ml urine sample, analyzed for presence such as ethanol, acetone, isopropanol, which indicates the use of alcohol or the presence of a potentially

pathological condition, such as diabetes mellitus. The method used was that of Goldbaum et al (15). Positives are determined by the gas liquid chromatography method, using the Hewlett-Packard 7620A module complex Research Chromatograph, for identification of peaks.

RESULTS AVAILABLE AND JUDGED MEANINGS

If a urine screening yielded a positive or positives, the urine sample subject was contacted by private means to ensure confidentiality and was then asked to voluntarily:

a. confirm the positive sample findings,

b. advise on voluntary use or clinically directed use of the substance identified,

c. state the reason for internal use of the substance identified,

d. state how the substance identified had been obtained,

e. state the frequency or situational use of substances identified.

If pathological significance was suggested from the urine sample, the subject was confidentially informed of the presence of the substance and its possible clinical significance.

Within the objectives and operational parameters and future applications of this study, positive urine screening results can be placed in a reciprocal relation to other sub-strategies findings and a judged meaning of significance can be made by the researcher.

CONCLUSION - TOXICOLOGICAL ANALYSIS

At the risk of over simplification, it may be stated that the drugs identified by urine screening in this study are primarily classed as depressants, drugs that diminish the state of alertness and decrease the impact the external environment has on the thoughts and feelings of the user. In spite of views strongly opposing drug use, the general acceptance of mood-modifying drugs by the general community is part of what by many is considered to be normal living habits (16).

In this study recommendations relate only to adaptation of established urine screening methods to the research design tested and no data is presented although the results were available. The results were most encouraging from the technical standpoint and application to future multistragegy correlation research are strongly indicated, and urgently suggested. Urine screening can be effectively used in the future phases of this study as follows:

1. determination of toxicological subject baseline

2. selection by screening of subjects for acceptance or rejection in a study

3. comparative performance of a subject, untreated and cleaned

4. subjects split on the basis of detected drug use

5. assessment of psychological and job behaviour through correlation between drug results and subjective questionnaire reporting.

BIOELECTRICAL ANALYSIS

MENTAL ACTIVITY INDEX

Heart rate variability, a research tool historically reserved for the medical sciences, affords methodological potential to the psychological and physiological sciences and to its application in the study of human performance as a field technique. Work reported (17, 18, 19, 20, 21) relates to its contribution to mental work and mental load. The obvious problem with this specific research strategy is its complexity of technology and the professional level of training required on the part of the researcher. Other problem areas lie in the standardization and interpretability of results as there are no absolutely clearly and conclusively defined standards against which judgement may be made (22).

STRATEGY OBJECTIVES

This study attempted an evaluation of standardization of recording methods and test conditions and aimed at determining the feasibility of drawing meaningful results from the data collected. While laboratory studies will be used in future studies for identification of critical biochemical variables related to load response in the operational environment, the heart rate field study hopefully will provide the linkage of theory with hypotheses and observations, and allow refinement of the research model.

This strategy rests upon the method of spectral analysis of sinus arrhythmia during mental loading. Heart rate irregularity will be used as a dependent variable during performance in an operational enviroment. The concept of sinus arrhythmia in this research strategy is not connected with pathology, i.e., irregularity increases under stress. A most interesting phenomenon is the suppression of the sinus arrhythmia (the irregularity of the heart rate pattern). There are indications that this occurs following changes in breathing pattern, a change in the vagal tone and that both kinds of suppression occur during mental load (23). It is now known that mental load will diminish irregularity without necessarily causing a rise in mean heart rate.

In future field studies, electrical heart activity as a dependent variable will be collected and analyzed. The most prominent electrical activity accompanying heart beat is the QRS complex in the ECG (the R-top). An approach to get the maximum amount of information out of an R-R time series will be formulated.

METHODS

Data collection was accomplished through remote transmitting. The subject was fitted with 2 ECG chest leads which were connected to a small 2 lead electronic transmitter carried on the body. The signal range was approximately 30 feet. A standard FM radio was used to receive the ECG signal on a FM band frequency and used for signal strength and drift monitoring. Through an amplifier the ECG signal was transmitted to a storage osciloscope which allowed further monitoring of quality. Accurate R-R counts in milli-seconds were transmitted to an instrument computer which is capable of storing the received pulse in appropriate, desired groupings. R-R counts were displayed on a second scope with the capability of recalling stored information for analysis. All ECG signals were recorded on tape for laboratory analysis.

RESEARCH INSTRUMENTS

Since only standard commercial equipment items were used in this study, no detailed operational description is offered, but instruments are identified.

For the recording and monitoring of continuous electrocardiograms during duty, the following equipment was taken to the field:

a. Skin cleaner for electrode application, a product of IVM, "In Vivo Metric" systems, P.O. Box 217, Redwood Valley, California 95470.

b. The electrodes for the two chest leads used, are a "Medi-Pad" disposable monitoring electrode, pre-gelled, manufactured by General Medical Corporation, Richmond, Virginia.

c. A small remote signal transmitter, pocket carried by the subject, transmitted on a discrete pre-selected FM frequency, manufactured by EKEG Electronics Co., Ltd., Vancouver, B.C.

d. A Sony, TFM - 7300 WA, AM-FM portable transistor radio to provide frequency monitoring by heart beat. The FM frequency range employed was 87.5 - 108 MHZ.

e. A portable, four channel, three speed instrumentation tape recorder using ¼ inch magnetic tape and capable of FM recording. The 3960 series recorder, manufactured by Hewlett-Packard Company, 690E Middlefield Rd., Mountain View, California 94040.

f. Recording tape used, Scotch, AV177 Tensor, low noise ¼ inch magnetic for heavy duty applications by Minnesota Mining and Manufacturing Company, St Paul, Minnesota 55101.

g. A display oscilloscope system consisting of display module and plug-in units. The power supply/amplifier module, the 5103M, the dual time base, a dual sweep plug-in. The system allows visual monitoring of ECG tracing quality and is manufactured by Tektronic Inc., P.O. Box 500, Beaverton, Oregon 97005.

Laboratory analysis of ECG data in this study required instrument items (e), (f), (g) in addition to:

h. A Tektronic Type RM 504 Oscilloscope, a low frequency, high sensitivity laboratory instrument providing accurate time and amplitude measurements.

i. A Fabri-Tek FT-1074 Signal Averager, an instrument computer providing memory, readout and arithmetic function in a main frame unit capable of accepting measurement control plug-in units. The unit can provide data to a data processor. The manufacturer is Fabri-Tec Instruments Inc., 5225 Verona Rd., Madison, Wisconsin 53711.

j. A PDP-9T a highly modified computer manufactured by Digital Equipment Corporation.

A copy of the analysis program used, appears as Appendix "B".

RESULTS - BIOELECTRICAL ANALYSIS

Within the parameters set for this sub-strategy for bioelectrical measurements, study objectives were met, in that it was determined that continuous monitoring of electrical cardiac activity of subjects employed in high workload operations in the field is feasible. Subsequent laboratory analysis of data is within the capabilities of the equipment used in this study.

Major points emerging are:

1. In operational studies involving psychological measurement of stress using QRS complex analysis, it will be necessary to develop a method of precisely determining work activity during the monitored period so that the activity and the physiological response can be analyzed. This could be a simple paper-pencil, pre-determined time intervals, activity scale scoring to be done by the subject. This would allow more meaningful correlation of findings.

2. The feasibility of graphic voice analysis (sound spectogram) as a potential correlate will have to be investigated. Voice recording during EKG monitoring was accomplished on an experimental basis in this study. Play back revealed marked voice quality and sentence velocity changes that could prove significant as a stress indicator.

3. It appears difficult to observe and measure experimentally physiological cost of performance or performance decrement because of the complexity of the ATS multi-dimensional task. It raises the possibility that we are not measuring what we should be measuring. Observation strongly suggest that a task experienced individual can and does compensate for any presumed performance decrement on a particular aspect of the overall task.

4. It is possible that the self reporting questionnaires sub-strategy applied in this test is a more reliable, more economical, and a high integrity method of determining performance cost.

SELF REPORTING QUESTIONNAIRES

ADEQUACY OF SLEEP AND MOOD

Conservation of sustained performance is a critical factor in operational effectiveness. Many variables have a direct effect on the ability of man to efficiently function in an operationally oriented sense. Some of these variables are:

a. selection,

b. motivation,

c. training,

d. the psycho-, and physio-physical quality of the work environment,

e. design of task specific equipment,

f. the nature of the task,

g. rest-work schedules.

Of these basic variables all of which are important in ultimately determining quality of performance related literature offers compelling evidence that quality and adequacy of sleep is possibly the single most fundamentally critical variable in the determination of the physiological cost of work (24). How much sleep man needs or how much sleep loss can be tolerated is extremely difficult to determine, and little is known about methods of economically and reliably measuring these parameters.

In this study subjective assessment strategies were used including self-report questionnaires obtained within the working environment rather than a laboratory setting. No attempt is made at performance decrement measurement due both to the sensitive nature of the task performed by the Air Traffic Controllers and possible ramifications of feedback of such measures to either the subjects or to their employers.

STRATEGY OBJECTIVES

In the design of the research model of this sub-strategy the general priority objective was to obtain reliable and valid indicators, within the stated parameters, of subjective assessments of sleep adequacy and quality as these might relate to future assessments of the physiological cost of performance. A second objective was to assess the validity, reliability and potential utility of existing questionnaires, while a third objective was the assessment of possible expenditures of time and money by the investigator to carry out a full scale study at a later date.

In the final analysis it is hoped that results will point the way to economical but effective future research strategies and that they will contribute to methodological advancement. As research instruments, three questionnaires were used:

1. Sleep log and Stanford Sleepiness Scale, provided in Appendix "C".

2. Mood Scale, provided in Appendix "D".

3. Goldberg General Health Questionnaire, provided in Appendix "E".

All three questionnaires are relatively simple to administer "paper and pencil" tests, having in the view of their authors, at least a mininal demonstrated utility and reliability. Each questionnaire is individually detailed in this Chapter.

SLEEP LOG - METHODS

The first measurement method employed is the Sleep Log, a simple self-reporting questionnaire in which the subject describes sleeping - waking behaviour on a 24 hour basis. The author and developers are Paul Naitoh, Ph.D. and Laverne Johnson, Ph.D., from the Naval Health Research Centre in California (25). Degree of sleepiness is assessed by the Standard Sleepiness Scale which appears on the reverse side of the questionnaire and was developed by Deurent et al (26).

In this research program subjects were requested to complete a questionnaire on the day following the duty day during which the field study occurred, and failing that, completion was requested following a duty day similar to the field study day in terms of shift worked and day of the week. In this manner methodological consistency was maintained. Similar instructions were given to both the ATS sample, and the DCIEM control group.

Although statistical testing was applied to the CYYZ (N=14) and CYKZ (N=9) sample groups as independent samples, the small size of the CYKZ sample made validity of statistical testing suspect, and the

CYYZ + CYKZ groups were combined into one (N=23) sample of a statistically adequate size for these preliminary efforts. The DCIEM control group was minimally acceptable with an N=16. The following groups were subjected to statistical testing in the Sleep Log category:

a. sample CYYZ, N=14

b. sample CYKZ, N=9

c. sample CYYZ/CYKZ, N=23

d. control DCIEM, N=16

e. NOK, N=23, (re-categorization is explained in the text)

f. OK, N=16 (re-categorization is explained in the text)

SLEEP LOG - SCORING SCHEME

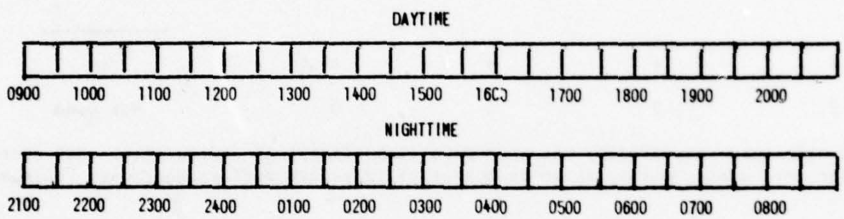
The author of this questionnaire, Dr. Paul Naitoh, was contacted regarding the use of the form and its scoring scheme. This investigator was authorized free use of the form and was advised to develop an independent scoring scheme. This was done and results were tabulated.

The scoring scheme determined for this study is applied to each question as follows:

1. Question #1.

<u>This question was not used in statistical testing.</u>

On the chart below draw a horizontal line through the squares corresponding to the half hour periods during which you were asleep during the last 24 hours. Put an X in the square corresponding to any half hour period during which you recall waking up for 15 to 30 minutes.



2. Question 2.

How much trouble did you have going to sleep last night?

[ ] None	[ ] Slight	[ ] Moderate	[ ] Considerable	Time to fall asleep ________ minutes
↑	↑	↑	↑	↑
0	1	2	3	Not scored

Scoring follows logically from the established idea that the most trouble free classification is represented by a zero score.

<u>The "time to fall asleep" value was not statistically tested.</u>

3. Question #3.

<u>This question was not used in statistical testing.</u>

How many times do you recall waking up last night?

________________ Not used.

4. Question #4.

How rested do you feel?

[ ] Well Rested	[ ] Moderately Rested	[ ] Slightly Rested	[ ] Not At All
↑	↑	↑	↑
0	1	2	3

Scoring is applied logically from 0 to 3, where zero is the trouble free classification, and the scoring is consistent with that of Question #2.

5. Question #5.

Do you feel that you could have used more sleep?

[ ] Yes	[ ] No
↑	↑
1 (NOK)	0 (OK)

This extremely significant question is a binary choice item rating the perceived requirement for additional sleep. The "yes" or "no" answer allows additional and truly different categorization to the traditional "sample" and "control" group, and appears ideally suited to this and future studies. Recategorization of the researchable universe of this study is introduced, and on the basis of this question, the additional groups are, and will be identified in all results as:

a. The "Not OK" group; NOK.

A "Yes" answer is scored as one and is identified in this study as NOK.

b. The "OK" group; OK

A "no" answer is scored as zero, the trouble free group, and is identified in this study as OK. This scoring is also consistent with the theme of scoring in Questions #2 and #4.

6. Question #6.

Today's Mood

[ ] Very Poor	[ ] Poor	[ ] Average	[ ] Good	Number of dreams recalled ________
↑	↑	↑	↑	↑
3	2	1	0	Not used

To maintain score consistency the good mood classifications scores zero. Any score different from the zero score, serves as an indicator of increased performance cost, "things are getting bad".

<u>The "Number of Dreams Recalled" data was not used.</u>

7. Question #7.

<u>This question was not used in statistical testing.</u>

Hours of work in last 24 hours?

__________________ Not used.

8. Question #8.

Choose one of the seven statements below which best describes your present feelings. How you feel right now.

0 → (1) Feeling active and vital; alert, wide awake.

1 → (2) Functioning at a high level, but not at peak; able to concentrate.

2 → (3) Relaxed; awake, responsive, but not at full alertness.

3 → (4) A little foggy; let down; not at peak.

4 → (5) Foggy; slowed down; beginning to lose interest in remaining awake.

5 → (6) Sleepy; woozy; prefer to be lying down; fighting sleep.

6 → (7) Almost in reverie; sleep onset soon; losing struggle to remain awake.

The Sanford Sleepiness Scale which includes 7 subjective feeling classifications is scored from zero to six following the same scoring principle established in the preceeding questions. Statement (1) is awarded a zero score indicating a most trouble free feeling, followed by a logical sequential scoring indicating greater negative feelings and an increased performance cost.

SLEEP LOG - ANALYSIS

In the first set of cases, in an attempt to establish the internal validity and reliability of independent scaling from one question to the next, a series of correlations were performed on the following comparisons:

a. question #2 with question #4

b. question #2 with question #6

c. question #2 with question #8

d. question #4 with question #6

e. question #4 with question #8

f. question #6 with question #8

In the Sleep Log results section, this analysis will be identified as "Case A".

The testing protocol for the correlational analysis is as follows:

1. The formula used for the test of statistical significance is:

$$F_{1,N-2} = \frac{r^2}{1-r^2}(N-2) \quad (27)$$

where r is the correlation of two items over the group or subgroup and is calculated by the formula

$$r = \frac{\Sigma XY - \Sigma X\Sigma Y/n}{\sqrt{[\Sigma X^2 - (\Sigma X)^2/n]\,[\Sigma Y^2 - (\Sigma Y)^2/n]}} \quad (28)$$

2. The level of significance is $\alpha = 0.5$

On the second set of cases, the Dunnett version of the t-test was performed on the differences between means of the sample (ATS, N=23) group and the control (DCIEM, N=16) group and between the OK Sleep Group (N=16) and the NOK Sleep Group (N=23) to test for the significance of observed differences in responses. The significance testing proceeds as follows:

1. ATS question #2 with DCIEM question #2

2. ATS question #4 with DCIEM question #4

3. ATS question #6 with DCIEM question #6

4. ATS question #8 with DCIEM question #8

5. OK Sleep Group with the NOK Sleep Group on Questions 2, 4, 6 and 8 respectively.

In the results, this analysis will be identified as "Case B". Test protocol is as follows:

1. The formula for Dunnett's Test is:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma 1^2}{N_1} + \frac{\sigma 2^2}{N_2}}}$$

(with $N_1 + N_2 - 2$d.f., (29) using Dunnett's Tables to correct for the effect of multiple t tests)

RESULTS - CASE A

Correlations were performed with two different sets of criteria:

1. The air traffic tower controller subjects as the sample group compared with the DCIEM subjects as the control group.

2. The re-categorization of the total study sample into an "OK" group and a "NOK" group on the basis of scoring of question #5 of the Sleep Log, as explained in the text.

Table T-1 shows the results of the correlations found for the combined CYYZ+CYKZ (N=23) sample. Table T-2 shows the results of the correlations found for the DCIEM group.

In Table T-1 of question #2 and #4 (a) yielded a non-significant correlation of r = .248. However, items (b), (c), (d) and (e) yielded correlations significant at the 5 percent level. The remaining correlation of question #6 with #8 (f) yielded a correlation of r = .363, which was non-significant.

TABLE T-1
COMBINED SAMPLE CORRELATION, CYYZ+CYKZ GROUP

	Items Compared	r	r^2	$1-r^2$	$F_{1,\ N-2}$	$F_{\alpha=@.05}$
a	2/4	.248	.060	.94	1.47	4.22
b	2/6	.483	.233	.77	6.30*	4.22
c	2/8	.608	.370	.63	12.39*	4.22
d	4/6	.426	.181	.82	4.62*	4.22
e	4/8	.582	.340	.66	10.92*	4.22
f	6/8	.363	.130	.87	3.15	4.22

* = significant at α = .05

TABLE T-2
CONTROL GROUP CORRELATION DCIEM GROUP

	Items Compared	r	r^2	$1-r^2$	$F_{1,\ N-2}$	$F_{\alpha=}$@ .05
a	2/4	.323	.1	.9	1.56	4.60
b	2/6	.269	.07	.93	1.05	4.60
c	2/8	.288	.08	.92	1.22	4.60
d	4/6	.542	.29	.71	5.72*	4.60
e	4/8	.827	.68	.32	29.75*	4.60
f	6/8	.338	.11	.89	1.73	4.60

* = significant at α = .05

Significant correlations of r = .542 and r = .827 were yielded on the comparison of questions #4 and #6, and #4 and #8 respectively. This is consistent with the findings on these two correlations in the CYYZ+CYKZ group.

Tables T-3, and T-4 show the results of the correlational analysis performed on the re-categorized "OK", (N=16), and the "NOK" (N=23) samples.

TABLE T-3
RE-CATEGORIZED SAMPLE CORRELATION, OK SLEEP GROUP

	Items Compared	r	r^2	$1-r^2$	$F_{1,\ N-2}$	$F_{\alpha=}$@ .05
a	2/4	.152	.02	.98	.28	4.60
b	2/6	.292	.09	.91	1.4	4.60
c	2/8	.114	.01	.98	.14	4.60
d	4/6	.590	.35	.65	7.56*	4.60
e	4/8	.529	.25	.75	5.18*	4.60
f	6/8	.236	.06	.94	.84	4.60

* = significant at α = .05

Only items (d) and (e) where in (d) question #4 was compared with question #6, and in (e) where question #4 was compared with question #8 were correlations yielded at significant levels of r = .590 and r = .529 respectively.

Only item (e) where question #4 and #8 were compared, did correlation occur at a significant level (r = .524). In all other correlation coefficients were non significant.

TABLE T-4
RE-CATEGORIZED SAMPLE CORRELATION, NOK

	Items Compared	r	r^2	$1 - r^2$	$F_{1, N-2}$	$F_{\alpha = .05}$
a	2/4	.042	0	1.0	0	4.32
b	2/6	.361	.13	.87	3.15	4.32
c	2/8	.389	.15	.85	3.78	4.32
d	4/6	.059	0	1.0	0	4.32
e	4/8	.524	.27	.73	7.77*	4.32
f	6/8	.012	0	1.0	0	4.32

* = significant at α = .05

RESULTS - CASE B

Table T-5 shows the results of Dunnett's Tests performed on the differences between the means of the scores on Sleep Log questions #2, #4, #6 and #8, for the CYYZ+CYKZ (N - 23) and DCIEM (N = 16).

Table T-6 shows the results of the t-tests performed on the differences between the means of the scores for the OK (N = 16) and NOK (N = 23) groups.

TABLE T-5
SAMPLE AND CONTROL T-TEST

CYYZ+CYKZ Group		N = 23	t calculated	Comparisons tested	t required for significant result	DCIEM Group		N = 16
#2	$\bar{X}$	.478	t=1.02	#2/#2	t=2.021	.625	$\bar{X}$	#2
	σ	.846				.957	σ	
#4	$\bar{X}$	.783	t=0.14	#4/#4	t=2.021	.813	$\bar{X}$	#4
	σ	.518				.834	σ	
#6	$\bar{X}$	.783	t=1.42	#6/#6	t=2.021	.5	$\bar{X}$	#6
	σ	.600				.516	σ	
#8	$\bar{X}$	1.435	t=.53	#8/#8	t=2.021	1.625	$\bar{X}$	#8
	σ	1.080				1.147	σ	

Note: all t calculated are non-significant

TABLE T-6
OK AND NOK T-TEST

OK Group		N = 16	t calculated	Comparisons tested	t required for significant result	NOK Group		N = 23
#2	$\bar{X}$	.125	t=2.86*	#2/#2	t=2.021	.826	$\bar{X}$	#2
	σ	.342				1.03	σ	
#4	$\bar{X}$	.313	t=5.79*	#4/#4	t=2.021	1.130	$\bar{X}$	#4
	σ	.479				.548	σ	
#6	$\bar{X}$	.375	t=2.86*	#6/#6	t=2.021	.869	$\bar{X}$	#6
	σ	.5				.548	σ	
#8	$\bar{X}$	.750	t=4.3*	#8/#8	t=2.021	2.04	$\bar{X}$	#8
	σ	.846				.928	σ	

* Significant at α = .05

In contrast to the T-5 result, in the OK and NOK analysis of variance tests, results on all questions, t-value obtained exceeded the critical value of t = 2.021.

CONCLUSIONS - SLEEP LOG

An overview of the statistical findings that flow from the analysis of the Sleep Log data demonstrate two facts that are important, both in terms of the present study, and for future applications of this assessment instrument. First, there is strong consistency shown in the relationship between several items of the Sleep Log. While the results do not show an absolute relationship, the demonstrated significant correlations show that an assessment of how well rested a person feels himself to be (question #4) is highly predictive of how positively he describes his present feelings (question #8). Similarly but with less certainty, one may infer that a parallel assessment of how well rested the individual feels (question #4) will be predictive of how good, or bad he assesses his mood to be (question #6).

There results are far from surprising or novel, but they are a preliminary objective demonstration of the folk lore fact that how well rested one is, determines to a great extent how good a person's mood is and how well prepared he is to perform his duties or work. The internal consistency established in this study argues for the further use of this paper and pencil test in future studies to examine the effect, on both individual responses and cluster of responses, of stressful environments and/or of the use of mood altering drugs and other tactics employed to counter stress.

Secondly, the absolute consistency of significant differences between the mean scores obtained on the Sleep Log items by the sleep OK and NOK groups, show that the four individual items tested (questions #2, #4, #6, and #8) are able to discriminate well between respondents reporting differently on the binary choice OK, NOK criterion. Again, this result indicates that the future use of this sub-categorization of groups on the binary choice basis will be a most profitable tool for the description and assessment of experimental and comparative groups as researchers attempt to establish the sources of stress and the responses of the individual both within and outside of his work cycle.

Taken together, the internal consistency and group to group significant differences strongly argue for the utility and applicability of this instrument to further studies of abnormal and stressful environments.

MOOD SCALE - METHODS

The Mood Scale allows the individual to describe subjectively his feelings of alertness, emotional state, social disposition and general mood. The scale is reputed to be a sensitive indicator of potential causes of performance decrement. The questionnaire was developed by the authors of the Sleep Log scales described earlier in this paper (30).

In the interest of consistency, the subjects participating were requested to complete the questionnaires, following the same instructions as those applied to the completion of the Sleep Log.

The groups tested were defined on the established criteria for this study as follows:

a. CYYZ + CYKZ (N=23) versus DCIEM (N=16)

b. OK (N=16) versus NOK (N=23)

MOOD SCALE - SCORING SCHEME

Scoring instructions, as these appeared in the literature, were followed and each of the four possible response categories were assigned a weight as follows:

NOT AT ALL	A LITTLE	QUITE A BIT	EXTREMELY
0	1	2	3

The scale contains 19 positive items representing the positive (P) score. These positive items reflect feelings and behaviour that generally decrease with inadequate or restless sleep. As a function of the weighting of the response categories, the P score can range from 0 (0x19=0), extremely sleepy, to 57 (3x19=57) extremely active and alert.

The scale contains a further 10 negative items representing the negative (N) score. Negative scores usually increase following sleep inadequacy and can range from 0 (0x10=0), extremely active and alert, to 30 (3x10=30), extremely sleepy.

The positive and negative items are identified on the Mood Scale questionnaire, as used in this study, as follows:

In addition to the N and P Mood Scale scoring schemes, for the purposes of analysis, a combined score labeled Composite (C), consisting of the (P) score minus the (N) score (P-N=C) has been calculated for each subject.

The derived indicators for analysis of the Mood Scale in this study are:

a. Sum of negative scores (N = 30 Max)

b. Sum of positive scores (P = 57 Max)

c. Sum of composite scores (C = P - N)

Scores N P C

A fourth parameter has been applied to the Mood Scale, as earlier stated in the methods section, which

uses the binary choice sub categorization item in the Sleep Log Scale. The significant feature of this categorization and scoring is sleep adequacy, which was defined as follows:

Do you feel that you could have used more sleep?

[ ] YES	[ ] NO
NOK	OK
↑	↑
1	0

NAME						AGE	SEX	DATE	
ITEM	NOT AT ALL	A LITTLE	QUITE A BIT	EXTREMELY	ITEM	NOT AT ALL	A LITTLE	QUITE A BIT	EXTREMELY
ACTIVE					GOOD-NATURED				
ALERT					GROUCHY				
ANNOYED					HAPPY				
CAREFREE					JITTERY				
CHEERFUL					KIND				
ABLE TO CONCENTRATE					LIVELY				
CONSIDERATE					PLEASANT				
DEFIANT					RELAXED				
DEPENDABLE					SATISFIED				
DROWSY					SLEEPY				
DULL					SLUGGISH				
EFFICIENT					TENSE				
FRIENDLY					ABLE TO THINK CLEARLY				
FULL OF PEP					TIRED				
SCORES:	N	P	C		ABLE TO WORK HARD				

MOOD SCALE - ANALYSIS

On the basis of the three scoring indices N, P and C, a series of Dunnett's tests were performed to determine if a significant difference existed between the means of the various groups. On the basis of the criteria established for the study, these groups are:

1. (CYYZ + CYKZ) versus DCIEM

2. OK versus NOK

The (CYYZ + CYKZ) represents the ATS tower sample group at N=23, while the DCIEM represents the control group at N=16. The second category identified the adequate sleep group, OK at N=16 and the non-adequate sleep group, NOK at N=23. It is noted that the similarity of the sample size in both groupings of the subjects is purely coincidental and does not suggest that NOK with N=23 is in fact the CYYZ/CYKZ sample at N=23. The basis for recategorization was sleep adequacy and not place of employment.

The testing protocol is as follows:

a. The assumption of a null hypothesis is that the (CYYZ + CYKZ) sample group represents an occupation not significantly different from the DCIEM control group in their responses to this questionnaire. The alternate hypothesis (H_1), states that a significant difference does exist between the groups tested.

b. Similarly the null hypothesis for the OK, NOK groups comparison assumes that these groups do not differ significantly on their responses while the alternate hypothesis states that the groups do indeed differ significantly.

c. The level of statistical significance is = .05.

d. Dunnett's test uses the formula:

$$t = \frac{\bar{X}_1 \quad \bar{X}_2}{\sqrt{\frac{\sigma 1^2}{N_1} \quad \frac{\sigma 2^2}{N_2}}} \quad \text{(with } N_1 + N_2 - 2 \text{ dif.)}$$

RESULTS - MOOD SCALE

Tables T-7 and T-8 show the results of Mood Scale Scores by the subjects as sample and control groups

and as re-categorized on the basis of sleep adequacy.

TABLE T-7
DUNNETT'S TEST FOR DIFFERENCE OF MEANS
(CYYZ + CYKZ) GROUP VS DCIEM GROUP

CYYZ+CYKZ Group		N = 23	t calculated	Comparisons tested	t required for significant result	DCIEM Group		N = 16
C	$\bar{X}$	26.13	t = .76	C/C	t = 2.021	28.81	$\bar{X}$	C
	σ	11.19				10.65	σ	
N	$\bar{X}$	4.74	t = 1.02	N/N	t = 2.021	3.5	$\bar{X}$	N
	σ	3.93				3.6	σ	
P	$\bar{X}$	30.74	t = 2.72*	P/P	t = 2.021	23.31	$\bar{X}$	P
	σ	8.60				8.26	σ	

* The results show that only in the positive category was there a difference between the scores at a significant value for the test at t = 2.72 (exceeding the critical value of t = 2.021).

TABLE T-8
DUNNETT'S TEST FOR DIFFERENCE OF MEANS
SLEEP OK GROUP VS SLEEP NOK GROUP

OK Group		N = 16	t calculated	Comparisons tested	t required for significant result	NOK Group		N = 23
C	$\bar{X}$	34.75	t = 4.06*	C/C	t = 2.29	.22	$\bar{X}$	C
	σ	11.00				7.29	σ	
N	$\bar{X}$	2.00	t = 3.5*	N/N	t = 2.29	5.78	$\bar{X}$	N
	σ	3.26				3.39	σ	
P	$\bar{X}$	36.75	t = 3.7*	P/P	t = 2.29	27.65	$\bar{X}$	P
	σ	8.39				6.17	σ	

* Significant at α : .05 level

When the groups were pooled and recategorized into sleep adequacy (OK) and non sleep adequacy (NOK) groups, table T-8 shows the results. Significant t values exceeding critical t values in all these indices were found.

CONCLUSIONS - MOOD SCALE

The results indicate that there are strong and significant differences in the Mood Scale, as a function of whether a subject see himself or herself as having experienced a good night's sleep. It is only when this sorting is applied, on the basis of subjective assessment of sleep adequacy that consistent differences are found across all three categories of scoring; N, P and C; indicating the potential utility of the Mood Scale as a sensitive and valid indicator of perceived subjective differences in mood following sleep. Again, indications are that the popularly held belief that sleep adequacy affects mood in a direct relationship, i.e., better mood relates to adequate sleep and poorer mood relates to inadequate sleep. In this case P scores were significantly higher for sleep OK than for sleep NOK and N scores were similarly lower.

For future use, this scale offers an opportunity to measure the individual's personal assessment of his mood and since it has been demonstrated here that this scale shows significant differences on the basic sleep OK, NOK categories, it would be expected that similarly finely tuned significant differences should be found for those using mood altering drugs or other tactics to cope with stressful environments.

GOLDBERG (GHQ) - METHODS

The 30 item, Goldberg General Health Questionnaire (GHQ) used in this study is a derivative of the long form, 190 questionnaire developed by Dr. David P. Goldberg (31). The 30 item questionnaire was made available by the Clarke Institute of Psychiatry for this study. No attempt was made to modify the general categories represented by the 30 items.

Presented in table T-9 is the questionnaire's layout.

TABLE T-9
THE 30 ITEM GHQ

CATEGORY	NUMBER OF ITEMS	LOCATION ON 30 ITEM GHQ
A General Health and Central Nervous System	1	Question #1
B Sleep and Wakefulness	4	Questions #2 → #5
D Observable - Personal Behaviour	4	Questions #6 → #9
E Observable - Relations with Others	2	Questions #10 → #11
F Subjective Feelings - Inadequacy, Tension Temper	10	Questions #12 → #21
G Subjective Feelings - Mainly Depression and Anxiety	9	Questions #22 → #30

In the interest of consistency, subjects were requested to complete this questionnaire following the same instructions as had been applied to all the other questionnaires used. The groups were derived on the basis of criteria established for this study as follows:

1. (CYYZ + CYKZ), (N=23) versus DCIEM (N=16)

2. OK (N=16) versus NOK (N=23)

GHQ (30) - SCORING SCHEME

The scoring method as suggested in the literature by Goldberg which he says offers diagnostic value and simplicity, is for example as follows:

Question #27 - been feeling reasonably happy, all things considered?

	more so than usual	about same as usual	less so than usual	much less than usual
	↑	↑	↑	↑
Score	0	0	1	1

This four-way scoring pattern forces the elimination of a central response tendency by assuring a left or right bimodal response choice.

It was felt that this scoring scheme could result in the loss of potentially valuable information, and in an attempt to clarify and perhaps improve the utility of the Goldberg scale, the content was examined logically and the decision was made to use a different, less mathematically discontinuous scoring scheme, as well as the dichotomous originator's preferred scheme. To this end, the following scoring was also applied:

Question #26 - been feeling hopeful about your own future?

	more so than usual	about same as usual	less so than usual	much less hopeful
	↑	↑	↑	↑
Score	-1	0	1	2

This scoring scheme is consistent with the approach established in this study, showing a numerical lowering of the score as subjective improvement occurs, suggesting better performance at decreased physiological cost. The first category listed is to be scored as -1 indicating improvement or movement away from the zero score, which itself identifies a "no change", "same as usual" situation. A score of 1 is indicating a worsening situation, while a score of 2 has a judgemental meaning of considerable negative potential.

The additional sample scoring parameter applied to the GHQ in this study is the binary choice item found in the structure of question #5 of the Sleep Log Scale.

GHQ (30) - ANALYSIS

For each of the scoring schemes applied (a) 0/0/1/1, and (b) -1/0/1/2, Dunnett's tests were performed

to determine if a significant difference could be demonstrated between the mean scores obtained. The groups once again are:

a. (CYYZ + CYKZ), (N=23) versus DCIEM, (N=16)

b. OK, (N=16) versus NOK, (N=23)

In the first group, the (CYYZ+CYKZ) group represents the total ATS tower sample and the DCIEM group is the control. The second group is a result of sorting of the entire population of this study (N=39) on the basis of the subjectively determined sleep adequacy thus creating the two groups identified.

The test protocol is established as follows:

1. The assumption is made that employment of the ATS group of subjects is not significantly more stressful than that of the essentially office and research oriented DCIEM control group. This *is the null hypothesis (Ho). The research hypothesis* (H_1) assumes a difference in either a positive or negative sense.

2. The statistical level of significance is $\alpha = .05$

3. The Dunnett's test formula used is:

$$t = \frac{\bar{X}_1 \quad \bar{X}_2}{\sqrt{\frac{\sigma 1^2}{N} \quad \frac{\sigma 2^2}{N_2}}} \quad (\text{with } N_1 + N_2 - 2 \text{ df.})$$

RESULTS - SAMPLE VS CONTROL

The results under the 0/0/1/1 scoring scheme are shown in table T-10, while table T-11 shows the analysis on the OK and NOK sleep adequacy re-categorization.

TABLE T-10
DUNNETT'S TEST OF DIFFERENCES OF MEANS ON THE GHQ USING THE 0/0/1/1 SCORING SCHEME (SCORE A) FOR THE ATS VS DCIEM GROUPS

CYYZ + CYKZ		N = 23	t calculated	Comparisons tested	t required for significant results	DCIEM		N = 16
Score Scheme 0/0/1/1	$\bar{X}$	2.78	t = .49	Score (a)	t = 2.021	2.19	$\bar{X}$	Score Scheme 0/0/1/1
	σ	4.44		Score (a)		3.15	σ	

TABLE T-11
DUNNETT'S TEST OF DIFFERENCES OF MEANS WITH THE 0/0/1/1 (SCORE B) SCORING SCHEME FOR THE OK VS NOK GROUPS

OK		N = 16	t calculated	Comparisons tested	t required for significant results	NOK		N = 23
Score 0/0/1/1	$\bar{X}$	.875	t = 2.69*	Score (b)	t = 2.-21	3.695	$\bar{X}$	Score 0/0/1/1
	σ	1.707		Score (b)		4.606	σ	

* The results indicate a significant difference only in the re-categorized sample analysis.

TABLE T-12
DUNNETT'S TEST OF DIFFERENCE OF MEANS ON THE GHQ
USING THE 1/0/1/2 SCORING SCHEME FOR THE ATS VS DCIEM GROUPS

CYYZ + CYKZ		N = 23	t calculated	Comparisons tested	t required for significant result	DCIEM		N = 16
Score -1/0/1/2	$\bar{X}$	-9	t = .81	Score (a)	t = 2.021	-11.25	$\bar{X}$	Score -1/0/1/2
	σ	10.18		Score (a)		7.11	σ	

TABLE T-13
DUNNETT'S TEST OF DIFFERENCES OF MEANS
WITH THE -1/0/1/2 SCORING SCHEME, SLEEP ADEQUACY GROUPS

OK		N = 16	t calculated	Comparisons tested	t required for significant results	NOK N = 23		
Score -1/0/1/2	$\bar{X}$	-13.93	t = 2.4*	Score (b)	t = 2.021	-7.13	$\bar{X}$	Score -1/0/1/2
	σ	6.92		Score (b)		9.38	σ	

* As with the Goldberg scoring scheme, the modified scoring scheme added in this study, yielded a significant difference only when applied to the re-categorized groups on the basis of sleep adequacy.

CONCLUSIONS - GENERAL HEALTH QUESTIONNAIRE

In parallel with the results obtained from the other paper and pencil tests, the GHQ results show that responses on this instrument do discriminate well between those reporting adequate sleep versus those reporting non-adequate sleep. This questionnaire extends the time frame of the self assessment back beyond the present or recent past in that it asks for the self evaluation over the past month.

In future studies it will be most useful in allowing the researcher to extend the time scale of his assessment of change and how it affects perceived general health, beyond the day by day assessments possible with the Sleep Log and Mood Scale.

CONCLUSIONS - COMBINED FINDINGS

What has emerged is that all of the applied scales are useful in that they allow a significant differentiation between people who consider themselves to be (a) sleep deprived or (b) not sleep deprived, in the sense only of adequacy of sleep as it was subjectively assessed. On that basis there is an internal consistency and validity to the Sleep Log in that solid correlations were demonstrated between how well rested the subject felt (question #4 Sleep Log) and how well he felt in terms of being ready and able to work (question #8) and to a slightly lesser extent with his present mood (question #6).

It is of interest that only one subject in the combined ATS group and again only one in the DCIEM group scored 4 (a somewhat negative response) in question #8 in the Sleep Log. All other subjects felt themselves to be in a relatively responsive and alert performance state. Essentially all scores in the Stanford Sleepiness Scale (question #8 Sleep Log) were in the 0, 1 and 2 range.

There is strong internal consistency in that, if the only information obtained concerned the adequacy of sleep (question #5 Sleep Log), reasonable predictions could be made about how well rested one would feel the next day and what the perceived mood would be. This means that some prediction about potential quality of performance can be made.

The judgement must be that the combined Sleep Log and Stanford Sleepiness Scale is useful. It offers excellent internal consistency and does identify items about individuals that should be known before measuring other variables. All items contained in the Sleep Log are seen to be valid subjective indicators because they do relate to each other logically and significantly.

The Mood Scale offers much useful information on the basis of the re-categorization of subjects into the subjective sleep adequacy NOK and OK groups. The questionnaire shows that significant differences exist between groups that have rated themselves as sleep adequate (OK) or sleep inadequate (NOK) on all three of the derived indicator scores (N, P and C) that emerge from the mood scale.

Findings on the basis of the division of the subjects into an (ATS) sample and (DCIEM) control group, indicate the only area where there is a significant difference is in the positive (P) index. This study strongly implies that the Mood Scale offers good internal consistency and reliability when applied on the basis of subjectives assessment of sleep adequacy, and thus is a useful tool in finer mood definition on a here and now basis in all of the three mood indices (N, P and C).

With the Goldberg questionnaire, in both the standard (0/0/1/1) binary scoring and the suggested modified (-1/0/1/2) scheme there appears to be no significant differences between the (ATS) sample group and the (DCIEM) control group. When the subjects are again re-categorized into OK and NOK groups on the basis of sleep adequacy, significant differences are demonstrated between their ratings. Thus a degree of interrelation between sleep adequacy, performance, mood and general health as subjectively determined is clearly apparent.

The conclusion must be that with the kind of agreement between the questionnaire used, the investigator should feel satisfied and encouraged in approaching larger groups on the basis of sleep adequacy categorization as well as job categorization. The possibility exists for continuing research on the further subcategorization on the basis of medication received to off set negative mood states and their effect of measured performance.

PERSONAL INTERVIEWS

WHY DO PEOPLE SELECT AIR TRAFFIC SERVICES

Very few of the controllers interviewed had selected the ATS field of employment as a first specific choice. Aviation interests were a guiding factor for most and for those who had obtained actual experience

in the piloting of aircraft, ATS was invariably a second option occupation. For the majority interviewed for whom a flying career was potentially available, the profession of airline pilot was discarded either because of fear of early medical loss of license, with the associated loss of earning power and security, or because the competition to reach the top as an airline pilot appeared too formidable an obstacle. This suggests that ATS employment was selected for reasons of achievement ease and long term security and because it was associated with aviation. For those who had never tried their hand at flying, it still held sufficient attraction to make many wish that they had at least given the flying of aircraft a try.

EMPLOYMENT SATISFACTION

Job satisfaction is of a high order in the group interviewed and during the personal interviews it became apparent that the level of satisfaction could well be substantially higher than the 88 percent indicated by the personal questionnaires. Typical comments echoed by many reflected this in statements such as:

"I like it 100 percent and would like to climb in the ATS hierarchy".

"Am really very happy. It's not a job, it's a pleasure".

"I look forward to coming to work".

"Would like to remain in the system".

Tower controllers come across as confident, strongly authority oriented, and critical of any influence negatively affecting their perceived high job status. In their subjective assessment of personal job quality they rigorously evaluate their own performance, and often compare their profession with that of the pilot, who represents the external human sub-system in the broad ATS man/machine systems complex.

Although there is the possibility that some of the comments expressed are influenced by the controllers suppressed desires to be pilots themselves, the comments do offer what must be considered strong evidence of the very positive view that tower controllers have of their professional performance. For example some said:

"You may be wrong, but you are never in doubt".

"Pilots have no more capability than we do: ATS is more demanding".

"I have not seen too many pilots who have the complete picture".

"As a pilot you control one aircraft - I control many".

"Some pilots have failed as controllers, controllers have never failed as pilots".

"I don't think pilots are superior. It does balance out, we are even".

"Being a pilot may be more glamorous, being an air traffic controller is more demanding".

"We are elite. We are the top of the heap".

There is little doubt that the work itself, that of controlling air traffic, met with an almost unanimously high degree of approval. It was particularly interesting that the tower controllers, without exception, rated moderate to high work load as favorable. They clearly preferred its challenge and responsibility and would regard with dislike low traffic density condition. When asked to subjectively assess their stress tolerance, the ratio of "high" and "above average" to "average" was 2:1. No one interviewed perceived his or her job stress tolerance to be less than average.

This does not suggest that the tower controllers were without criticism of either the system or their peers and profound dissatisfaction was expressed concerning policy, administration and working conditions. One feature of the dislike of management was that the further away either geographically or professionally, the more serious the criticism became, that is, local supervision was accepted, regional supervision was suspect and Ottawa based supervision was regarded with profound concern. It was generally felt that in the decision making process, very little participation was allowed or requested from the operational level, the actual "sharp end" of the ATS system. There was a high degree of agreement that any visible controller participation in management carried little weight and this resulted in feelings of abandonment, aloneness and deep frustration. The "management/labour" relations which claim to be attempting to help the controllers were very much considered a major area for producing high stress.

THE WORKING ENVIRONMENT

In this study no attempt was made to evaluate the technological status and job related efficiency of the electronic equipment used in controlling of air traffic. However opinions were solicited and observation were made that reflect on the professional and personal habitability of the physical tower cab, even if we assume equipment suitability. Generally, controller satisfaction and acceptance was low. Comments that reflected the dislike were for example:

"The place is very depressing, so I dress brightly".

"When overcast, the place is deadly".

"I do not like its remoteness. It should be where the people are".

"Glare is disturbing".

"The tower is a tomb".

"Climatic control and layout are poor".

"The tower is full of flies and internal window cleaning is a ridiculous spectacle".

"No restroom facilities and dirty locker rooms".

It is quite surprising that tower cab design has incorporated so little of available information and knowledge about design of work space. This investigator feels that one of the reasons that more strongly voiced objections to tower cab design have not been raised is that tower controllers who must look out, can look out and have a wide, often fascinating and colourful panorama, filled with motion and variety and they attend to that rather than the unpleasantness in which they are ensconced. It is a visual panorama that represents the focus of their responsibility. As a working space it remains unsatisfactory and by the concern it causes and the adverse effect it has on functional effectiveness, it must be assessed as a contributory stress factor, increasing physiological performance cost. Habitability was clearly compromised by each of the negative factors raised by the tower controllers.

SOCIAL NEEDS AND EGO

Man's need to belong to a group, no matter how vaguely defined or ethereal is well satisfied by being an air traffic controller. The group under observation had, by its own assessment and design, established itself well as a strongly elite entity. Attempting to fit in has consequences not always to the advantage of the incoming controller. Because some had been unable to comply in specific areas of perceived elite demands, the formation of small sub-groups had resulted. These were ill defined but capable of accommodating specific individuals that have to be brought into the fold so peer pressure was thus applied.

Self image is strong and positive and is reflected in the mode of dress and general grooming which, in the facilities observed in this study, is at a very high level. It is clear that punctuality and performance, in addition to immaculate appearance, are strong ego feeding factors and are of paramount importance to the controllers. They add, in considerable measure, to improved self-image and confidence. All this translates into making ATS controllers well motivated achievers.

Management would do well to recognize this and realize that proper usage of ego boosting ploys and paying attention to the areas of personal dignity and pride could well be the most powerful tool in management armamentarium. The tower controllers who were observed think of themselves as the best in the world. Quite possibly they are.

PLACE OF RESIDENCE AND RECREATION

In the selection of the location of a place of residence the air traffic controllers were practical and made a decision on the basis of (a) type of dwelling, (b) social class of others living in the area and (c) proximity of resources such as shopping centres, transportation and educational facilities in that order of priority. The majority felt that a significant feature of their dwelling should be that it have an area such as a den that would provide privacy and a place to "unwind". Decoration would be subdued in subtly blending warm, dark tones. The uniformily expressed need for this privacy suggested the major importance for this requirement.

Their recreational needs were predominantly non-team, non-violence oriented and reflected the preference for out-door open space activities and quiet indoor hobbies with a high degree of creativity and personal satisfaction. Not surprisingly in view of the stringent medical standards for the profession, the majority of the controllers related strongly to the need for physical fitness. In these needs and hopes they did not differ significantly from the control group.

CONCLUSIONS - INTERVIEWS

The objective is to make an informal appraisal of how ATS tower controllers respond to their environment, what their subjective assessment of ATS is and how these factors might relate to physiological performance cost. Unfortunately, as in any other area of mythology, a body of non-scientific attitudes, opinions and ideas have evolved around the role and being of the air traffic controller, giving ATS an often mystical and mythical status. While some of the myths contain elements of truth and reality, occasionally even including factual accounts of their origins, there is an overwhelming feeling that one is dealing not with scientific studies but with a collection of myths.

With the concurrence of those being studied, an attempt has been made in this appraisal to let the tower controllers give their side of the story as well as formulating analysis on the basis of unobtrusive observation of activities over a specified period of time. A better understanding of who controllers are and what their approach to life and work is, is essential to the direction future research must take. So to this end, this sub-strategy allowed the controller to "free-associate" during the personal interview, so that a better insight into the ATS human could be obtained.

What emerged is that one must conclude that we are dealing with highly motivated professional individuals who have felt and manifestly responded to the need to become something that is worthwhile. They are involved in a type of employment that forces them to mature rapidly and one in which a high degree of satisfaction is experienced, probably more so than is the case with most people who work. It appears that air traffic controllers have much in common with people in other occupations and do not appear to differ signifiwhen specific comparisons are made.

But because of the nature of their work, they have very special needs and it is suggested that at higher levels of management and especially under the levelling influence of the Public Service, these needs are not easily recognized nor responded to when uncovered. This is perhaps so in part because viable answers

to potentially difficult problems are not available. There is a strong underlying theme brought to light for the author of this analysis that research programs in all their sophistication may be grossly misdirected, ignoring in their recommendations many of the subtle yet significant clues that can be gleaned from in-depth interviews and observations of and with operating controllers. The impression given when one compares published research and pseudo-research findings in this field with the often critical and caustic comments of the controllers is that a series of bad or faulty starts taken at the early stages of research into the stressful nature of the controllers environment have been pursued as if they had been landmarks in the research into the problem rather than the uncertain efforts that they often are. On some inaccurate assumptions and questionable research, the plight of the controller has in part been misconceived, misviewed and the resolution of the problem grossly mismanaged.

STUDY SUMMARY

There appears to be general agreement in the research literature (32) that when an individual admits to inadequacy of sleep, there will be a tendency on his part to react uneconomically to external stress. Work performance often remains intact, which suggests that subjects have a capacity to cope with the anxiety causing situation and this may possibly be a function of their span of energy. This suggests that under such circumstances, adaptation occurs at the behavioural level and at the level of physiological arousal (ref 33, 34) but there is serious potential for adverse effects as the subject is investing considerable effort to achieve a performance standard.

It is the main objective of this paper to demonstrate the probability that excessive physiologic cost expenditures occur in stressful environments and to test instruments allowing measurements relevant to this issue to be made in the actual working environment, with minimum job interference to the subject. The Air Traffic Services environment was selected first because it is assessed as being highly stressful, and second to encourage increased research effort over what has thus far been expended on Canadian ATS studies. The research model is designed in a manner that allows input sample substitution so the methods and instruments may be profitably employed in other stressful work environments.

The statistical analysis aspect of the study indicates that the independent ATS tower groups of CYYZ (N=14) and CYKZ (N=9) are too small and too highly variable to allow the making of predictive statements with any degree of certainty of the significance of any difference measured between these two groups or between each and the DCIEM comparison group.

In this study results confirm the utility, reliability and validity of the questionnaires and indicate even greater potential power when used with further sub-categorization of the sample on the basis of sleep adequacy, rather than on the basis of location of employment.

In addition to the questionnaires which were scored and analysed, there were three other sub-strategies tested that provided a basic dress rehearsal of the total systems research model of this study pertinent to the investigation of the effect that the external environment has on air traffic tower controllers. These sub-strategies are:

a. toxiological urinalysis screening,

b. continuous monitoring of electrical cardiac activity, and

c. behavioural observation - interview results.

The three ((a), (b) and (c)) sub-strategies were not intended to yield data of useable statistical nature for this preliminary study and did not allow analysis within the framework of this study. But there is compelling evidence from the literature that each of these strategies may be a significant contributory factor in future research in that the urine screening for example suggests further sub-categorization of the sample into drugs, and non-drug users, specifically in terms of mood-modifying substances. This would make possible investigation of the effects drug use has on a person's work performance, on their mood as measured by the Mood Scale with its demonstrated consistency and reliability, or on other indicators of physiological cost, such as the ECG/QRS activity assessments. Urine screening satisfies a critical research parameter for future studies in that the presence of mood-modifying drugs is a confounding factor when measuring the effects that any alteration of the environment may have on levels of stress and levels of performance. The effect of the presence of mood-modifying drugs must be separated from the effects of external environmental change, and this strategy allows that by indicating the presence of such drugs.

The high cost, high technology ECG/QRS activity research strategy provides a method of determination of baseline QRS values which will be required at the commencement of any study involving the measurement of any effect that altering the environment may have on mental activity as compared to performance in the unaltered environment. Transit and terminal QRS scores would then become mental activity change indicators. Urine screening, another high cost, high technology strategy, would allow detection of the presence of mood-modifying drugs which if they were introduced at a mid-point of the long term study would invalidate or at least alter data resulting in, at best, suspect findings.

Emerging from all of this is that, potentially, it may well be that the easily administered and very economical paper and pencil research methods do provide reasonably good indicators that parallel the physiological indicators. Once, and if, this is affirmed, it could well be that in the final research stage the expensive toxicological and bioelectrical methods could be eliminated in favor of the paper and pencil strategy of mood, general health and sleep quality assessments.

The behavioural observations and interviews are significant for future studies in that they provide a research instrument allowing both objective and subjective assessment of behavioural changes. A new and separate field of behavioural psychology has emerged and is based on the tremendous power and influence on behaviour of changing perceptions and expectations. One of the strongest methods of changing the internal state of the organism is to change its perception of an altered environment, in that if you change the

environment, and the changes are imperceptable (sub-perception threshold), it may well be that the subject perpetuates in assuming that things were as they have always been, and thus continues to function in its pre-change stressful way of responding in what should have become a non-stressful situation. By the same token, it is conceivable that the perception of an environment can be modified (without actually altering that environment) in such a manner that those in the situation will begin to function in a less stressful way, in response to an unchanged environment.

This may well be the most effective if not the only effective intervention method available in response to some stressful environments. Scientific investigation of the human response patterns in the ATS for example could be studied by graphic voice pattern analysis (sound spectogram), and the measuring of job performance quality. Thus the behavioural assessment research strategies, both objective and subjective must be an integral part of any multi-strategy model concerned with physiological performance cost.

FUTURE RESEARCH

This study established a research methodology possessing sufficient flexibility and integrity to cope with growth contingencies. On the basis of the findings a future research systems model evolved and is shown in Fig. 2. With regards to the flexibility of the research design proposed, this study demonstrated the need to consider constantly the growth potential, and to apply a detailed feasibility analysis as follows:

1. Restraint analysis.

   Cost limitations and time constraints are factors the design must be able to absorb, since these may well appear when the program is fully active, but not yet completed.

2. Technological feasibility analysis.

   Methodology must be planned on the basis of current and developing states of the art in technology since the integrity of any research program is strongly dependent on available technology but must not be allowed to be limited by out-dated or less than fully developed technical research strategies or instruments.

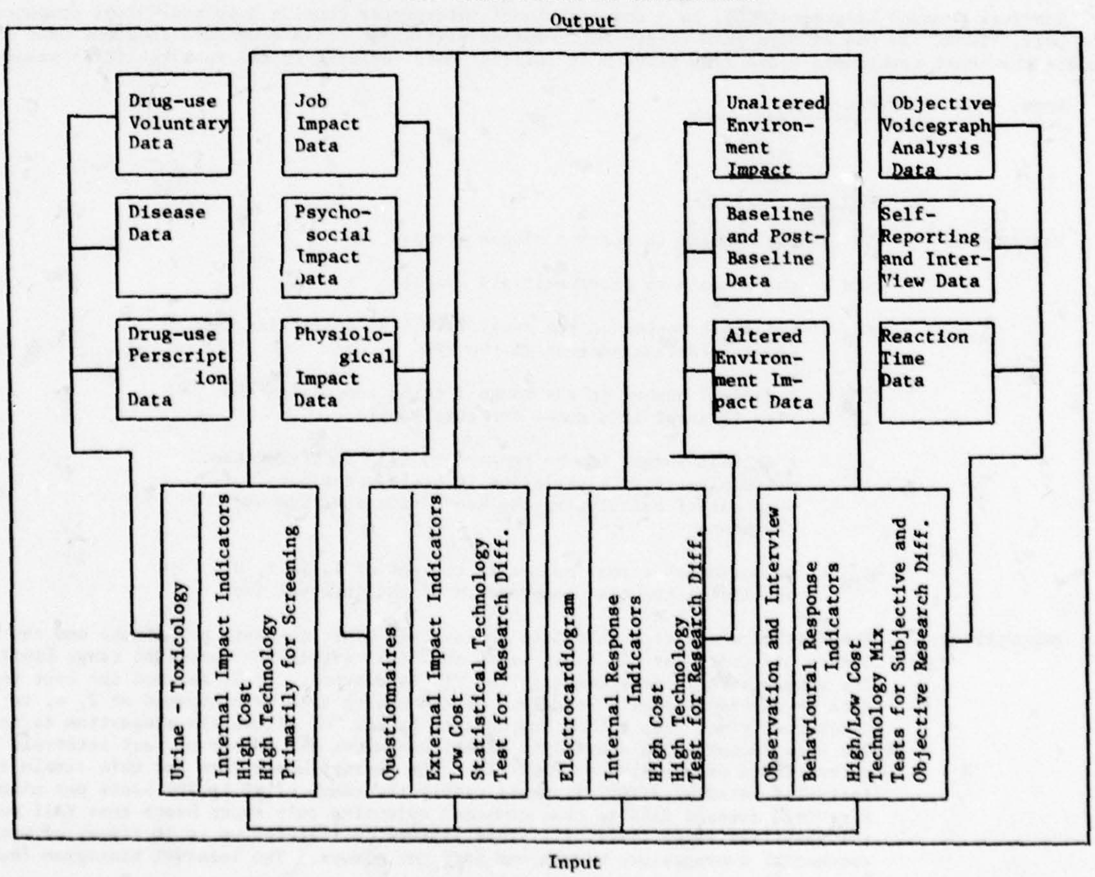
3. Growth analysis.

   Growth potential must be designed into the system so that in case of available mid-program advanced technology, no termination or costly re-design with associated time loss, results.

Real world factors often distort theoretically achieved design flows. Part of any planning process must be to prevent if possible or at least to minimize the impact of such occurrences.

FIGURE 3

RESEARCH SYSTEM DESIGN FOR SUBSEQUENT STUDIES OF ATS AND OTHER STRESSFUL ENVIRONMENTS



APPENDIX A

METHOD FOR LIBRIUM AND VALIUM (CHLORDIAZEPOXIDE AND DIAZEPAM)

1. 1 ml Urine in 15 ml screw cap tube. Add 0.1 ml Con. HCl.

2. Place tube in oven at 100°C for 2 hours.

3. Cool tube and add 0.4 ml NH_4OH.

4. Add 3 ml Chloroform, shake for 15 mins. Centrifuge and transfer organic phase to another 15 ml conical tube.

5. Evaporate to dryness. Residue dissolved in 3 drops MeOH and apply to thin layer plate.

6. Develop plate in Benzene up to 10 cm.

7. Dry plate and spray with:

   Reagent #1. (2% freshly prepared Sodium Nitrite in 1N HCl), dry and spray with Reagent #2. (2% aqueous solution of N-1 Naphthylethylenediamine dihydrochloride).

INTERPRETATION

Blue spot at Rf. .44 corresponding to ACB indicates usage of either of the benzodiazephine.

Yellow spot 1-Methyl-2 amino-5 chlorobenzopheno (MACE) at Rf .70 before spraying any reagent, indicates very recent usage of diazepam.

Flurazapam gives a blue spot slightly higher than ACB.

10 mg dose of either diazepam/chlordiazepoxide will give positive reaction for ACB for 5 days.

Dr. B.M. Kaput
Director of Laboratories
Clinical Institute
Addiction Research Foundation

20 November, 1974.

APPENDIX B

ANALYSIS PROGRAM - BIOELECTRIC MEASUREMENTS

Fabritek Command Language (FCL), is a conversational interpreter running into the PDP-9T computer. To the user, "SCORE" is one of more than twenty FCL commands available. This command allows the user to calculate the heart arrhythmia score from heart beat interval data residing in the Fabritek (FTK) memory.

FORM: > Score A T L C

or

> SCM A T L C

VARIABLES:

SC ; the command to score a single sample

SCM ; the command to score multiple samples

A_{10} ; a decimal number in the range Ø to 4095 specifying the sample starting address in the FTK

T_{10} ; a decimal number in the range 1 to 63 specifying the time interval in minutes for this sample

L_{10} ; a decimal number in the range 1 to 127 specifying the maximum heart rate deviation in beats per minute for purposes of calculating the heart rate coverage and the score

C_{10} ; an optional decimal number in the set of 1, 2, 4, 8 specifying the time compression of the interval data

DESCRIPTION: The SCORE command will calculate the average heart in beats per minute and the heart arrhythmia score for all heart beats that fall within the specified range limit "L". The sample size is determined by the "T" is minutes. It is assumed the beat interval data is recorded in milliseconds. If the analog data is converted at 2, 4, or 8 times real time, then a time compression factor "C" allows the alogarithm to account for this factor. The alogarithm first determines the number of beat intervals required for a sample size of time "T". The average heart rate for this sample is initially calculated for all beats within the range of 40 to 200 beats per minute. A revised average rate is then computed rejecting only those beats that fall outside the specified range limit "L". This process is iterated up to 10 times, or until sequential averages are within one beat per minute. The interval histogram count is

printed following the SCORE. The number of beat intervals exceeding "L" is also printed, followed by the FTK address of the last beat interval comprising this sample.

APPENDIX C
SLEEP LOG - QUESTIONNAIRE

NAME

1. On the chart below draw a horizontal line through the squares corresponding to the half hour periods during which you were asleep during the last 24 hours. Put an X in the squares corresponding to any half hour period during which you recall waking up for 15 to 30 minutes.

DAYTIME

0900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 2000

NIGHTTIME

2100 2200 2300 2400 0100 0200 0300 0400 0500 0600 0700 0800

2. How much trouble did you have going to sleep last night?
( ) NONE ( ) SLIGHT ( ) MODERATE ( ) CONSIDERABLE

Time to fall asleep ________ MINUTES

3. How many times do you recall waking up last night? ________

4. How rested do you feel?
( ) WELL RESTED ( ) MODERATELY RESTED ( ) SLIGHTLY RESTED ( ) NOT AT ALL

5. Do you feel that you could have used more sleep? ( ) YES ( ) NO

6. Today's Mood
( ) VERY POOR ( ) POOR ( ) AVERAGE ( ) GOOD

Number of dreams recalled

REMARKS (Note especially reasons for loss of sleep, such as duty, noise, cold, personal, etc.

SLEEP LOG

8. Choose one of the seven statements below which best describes your present feelings. How you feel right now.

(1) Feeling active and vital; alert, wide awake.

(2) Functioning at a high level, but not at peak; able to concentrate

(3) Relaxed; awake, responsive, but not at full alertness.

(4) A little foggy; let down; not at peak.

(5) Foggy; slowed down; beginning to lose interest in remaining awake.

(6) Sleepy; woozy; prefer to be lying down; fighting sleep.

(7) Almost in Reverie; sleep onset soon; losing struggle to remain awake.

APPENDIX D
MOOD SCALE - QUESTIONNAIRE

Instructions: For each item, choose one of the four answers that best describes how you feel now. Then put an "X" in that box.

NAME						AGE	SEX	DATE	
	NOT AT ALL	A LITTLE	QUITE A BIT	EXTREMELY		NOT AT ALL	A LITTLE	QUITE A BIT	EXTREMELY
ACTIVE					GOOD-NATURED				
ALERT					GROUCHY				
ANNOYED					HAPPY				
CAREFREE					JITTERY				
CHEERFUL					KIND				
ABLE TO CONCENTRATE					LIVELY				
CONSIDERATE					PLEASANT				
DEFIANT					RELAXED				
DEPENDABLE					SATISFIED				
DROWSY					SLEEPY				
DULL					SLUGGISH				
EFFICIENT					TENSE				
FRIENDLY					ABLE TO THINK CLEARLY				
FULL OF PEP					TIRED				
					ABLE TO WORK HARD				

APPENDIX E
GOLDBERG GENERAL HEALTH QUESTIONNAIRE

30 ITEM U.S. SCALE

IN THE PAST MONTH:

1. - been able to concentrate on whatever you're doing?

better than usual	same as usual	less than usual	much less than usual

2. - lost much sleep over worry?

not at all	no more than usual	rather more than usual	much more than usual

3. - been feeling mentally alert and wide awake?

better than usual	same as usual	less alert than usual	much less alert

4. - been feeling full of energy?

better than usual	same as usual	less energy than usual	much less energetic

5. - been having restless, disturbed nights?

not at all	no more than usual	rather more than usual	much more than usual

6. - been managing to keep yourself busy and occupied?

more so than usual	same as usual	rather less than usual	much less than usual

7. - been getting out of the house as much as usual?

more than usual	same as usual	less than usual	much less than usual

8. - been managing as well as most people would in your place?

better than most	about the same	rather less well	much less well

9. - felt on the whole you were doing things well?

better than usual	about the same	less well than usual	much less well

10. - been able to feel warmth and affection for those near to you?

better than usual	about same as usual	less well than usual	much less well

HAVE YOU RECENTLY:

11. - been finding it easy to get on with other people?

better than usual	about same as usual	less well than usual	much less well

12. - felt that you are playing a useful part in things?

more so than usual	same as usual	less useful than usual	much less useful

13. - felt capable of making decisions about things?

more so than usual	same as usual	less so than usual	much less capable

14. - felt constantly under strain?

not at all	no more than usual	rather more than usual	much more than usual

15. - felt you couldn't overcome your difficulties?

not at	no more	rather more	much more
all	than usual	than usual	than usual

16. - been finding life a struggle all the time?

not at	no more	rather more	much more
all	than usual	than usual	than usual

17. - been able to enjoy your normal day-to-day activities?

more so	same as	less so	much less
than usual	usual	than usual	than usual

18. - been taking things hard?

not at	no more	rather more	much more
all	than usual	than usual	than usual

19. - been getting scared or panicky for no good reason?

not at	no more	rather more	much more
all	than usual	than usual	than usual

20. - been able to face up to your problems?

more so	same as	less able	much less
than usual	usual	than usual	able

21. - found everything too much for you?

not at	no more	rather more	much more
all	than usual	than usual	than usual

22. - been feeling unhappy and depressed?

not at	no more	rather more	much more
all	than usual	than usual	than usual

23. - been losing confidence in yourself?

not at	no more	rather more	much more
all	than usual	than usual	than usual

24. - been thinking of yourself as a worthless person?

not at	no more	rather more	much more
all	than usual	than usual	than usual

25. - felt that life is entirely hopeless?

not at	no more	rather more	much more
all	than usual	than usual	than usual

26. - been feeling hopeful about your own future?

more so	about same	less so	much less
than usual	as usual	than usual	hopeful

27. - been feeling reasonably happy, all things considered?

more so	about same	less so	much less
than usual	as usual	than usual	than usual

28. - been feeling nervous and strung-up?

not at	no more	rather more	much more
all	than usual	than usual	than usual

29. - felt that life isn't worth living?

not at	no more	rather more	much more
all	than usual	than usual	than usual

30. - found that you couldn't do anything because your nerves were too bad?

not at	no more	rather more	much more
all	than usual	than usual	than usual

APPENDIX F
PERSONAL INFORMATION QUESTIONNAIRE
1975 - ATS - 1976

Please answer the following questions by circling one response or filling in the blank space. There are no trick or right or wrong answers, so please respond the way you really feel. Complete confidentiality is assured once the form has been completed. Information will not be used in any manner that allows identification, and is requested in support of a study concerning physiological cost determination during continuous activity.

Your participation and support are sincerely appreciated.

---

1. Name:
2. Sex: Male/Female
3. Age:
4. Marital Status: Single Married Other*

   Separated Divorced

   * (if "other", please elaborate)
5. Any children:
6. Education:
7. How long have you been associated as an Air Traffic Controller:

   1 - 5 years 6 - 10 years 10 - 15 years

   16 - 20 years 20 - 25 years 26 - up
8. How did you become interested in ATS?
9. Did you have any previous occupation?
   (if affirmative, please elaborate).
10. Are you satisfied and happy with the actual job activity?
    (would you rather fight than switch).
11. Are you a pilot? (Do you hold a valid pilot license?)
12. Are you receiving pilot training? (Do you hold a valid student pilot permit?)
13. Would you like to become a pilot?
14. Is flying your "first love"?
15. In your own estimation how do you consider your stress tolerance?

    Low Average Above average High

    (stress is any external event that causes a reaction of displeasure or anxiety in you)
16. Do you have a "code of ethics" by which you live?

    (any of the ten commandments or any non-religious standard to which you have decided you must conform - please elaborate).
17. Are there ATS facilities that have an "Elite Status" among all?
    (if it makes you feel more comfortable, use the U.S. facilities as an example).
18. If you answer to 17 is affirmative, would you state how "Elite Status" is achieved.
19. In selecting your place of residence, would you make your choice on the basis of:

    a. the class of people living in that neighbourhood

    b. the type of dwelling

    c. availability of resources (shopping plazas, parks, etc.)

    d. other (please elaborate)
20. What sports are you actively participating in?
21. What are your hobbies?
22. In your estimation, what kind of person has promising ATC potential? (consider character, personality,

education, desire, age limits, attitude towards life, etc.)

23. Is there anything you would like to say with regard to the questions asked or subject omitted? Please take your time in answering. It may be a thought about yourself, your work or your friends.

APPENDIX G
PERSONAL INTERVIEW FORMAT

Guide lines used in personal interviewing of ATS personnel. The information requirement are specific to this study only.

1. Concerning the internal tower environment;

   a. how do you feel about the colour schemes

   b. does working your position in sunlight affects you differently to when it is in the shade, and what is your preference

   c. does the tower design, and the equipment layout, please you professionally

   d. do visits to the Tower by individuals or groups distract you

2. Concerning the external tower environment;

   a. is parking lot accessibility good

   b. are refreshments and foods facilities satisfactory

   c. is there an area allowing you to relax or do you feel you need such a place

   d. are washroom facilities readily available and adequate

3. Concerning your peer associations during duty; (question to male);

   a. do you prefer an all male to a mixed male/female group

   b. does the presence of a female inhibit you in any way

   c. do you release professional frustrations in the presence of a female or do you delay reactions

   d. do you actually prefer the presence of a female controller

4. Concerning your associations during duty (questions to females);

   a. does being a minority bother you in this professional setting

   b. are you aware of any effect you cause in your male peers

   c. would you welcome a protective attitude towards you

   d. do you allow strongly directed feelings non-professional in nature, and do these interfere with your work

5. Concerning territoriality of your operating position;

   a. do you feel strongly protective of your position

   b. do you allow interference and if so, to what extent

   c. does your position become a strongly private place

   d. does stress in any way affect your attitude towards positional territoriality

6. Concerning flying interests:

   a. are you now a pilot

   b. how far back did you desire to fly

   c. what interferred with your flying aspirations

   d. would you rather fly than do your present job

7. Concerning your attitude towards those who fly;

   a. do you feel antagonistic towards pilots

   b. in your duties, do you control aircraft (impersonal) or do you control pilots (personal)

c. are you superior to a pilot professionally

d. does a pilot have your respect or do you feel more like his servant

8. Concerning your employment in ATS;

a. is job isolation (perhaps through the use of a head set) important to you

b. is working space a factor, in that you would prefer compactness through positional isolation (not unlike an aircraft cockpit)

c. would employment in the centre (a large space) be attractive to you

d. is direct supervision objectionable to you

9. Concerning the making of professional errors;

a. does it change the significance if others are present (bothers you much more or much less)

b. does it change the significance of the error, if others are aware of it

c. do you expect or desire your peers to intervene

d. does having made an incorrect professional decision bother you for a long time

10. Concerning external recreational activities;

a. you participate in any high risk activity (parachuting, hang gliding, car racing, etc.)

b. are your sports activities indoors or outdoors oriented

c. do you prefer team sports

d. does isolated camping in "unspoiled" nature appeal to you profoundly

REFERENCES

1. Maxwell, V.B.: "Stress in Air Traffic Control". Paper presented at the University of Manchester, October, 1973.

2. British Medical Journal: "Happy Landings", article on ATS responsibility for air safety, 421, August, 1973.

3. Cobb, S. and Rose, R.M.: "Hypertension, Peptic Ulcer and Diabetes in Air Traffic Controllers", Journal of the American Medical Association, Vol. 224, 1973, pp 489-492.

4. International Labour Office: "Conditions of Employment and Service of Air Traffic Controllers", Geneva, Switzerland, 1972.

5. Selye, H.: "Stress without Distress", McClelland and Stewart, Toronto, 1974.

6. Wagstaff, A.E.: "Clinical Problems and Stress in Air Traffic Control", The Controller, 1974, 13(2): pp 7-8.

7. Wilson J.G.: "Fatigue and the Controller", The Controller, 1972, 11 (1-4): pp 7-8.

8. Darley, J.T.: "Management Factors in Reducing ATCS Stress", The Controller, 1969, April, pp 13-15.

9. Melton, C.E.: "Comparison of OPA LOCKA Tower with other ATC Facilities by Means of a Biochemical Stress Index", FAA - AM - 74 - 11.

10. Mathews, J.J.: "Attitudes on en Route Air Traffic Control Training and Work: A Comparison of Recruits Initially Trained at the FAA Academy and Recruits Initially Trained at Assigned Centers", FAA - AM - 75 - 3.

11. Lalonde, M.: "A New Perspective on the Health of Canadians: A Working Document", Government of Canada.

12. Sunshine, I.: "Manual of Analytical Toxicology", The Chemical Rubber Co., Cleveland, Ohio, 44128.

13. Meola, J.M. and Vanco, M.: "Use of Charcoal to Concentrate Drugs from Urine Before Drug Analysis", Clinical Chemistry, Vol. 20, No. 2, 1974.

14. Stahl, E.: "Thin - Layer Chromatography", A Laboratory Handbook, Second Edition, Springer - Verlag, New York, 1969.

15. Goldbaum, L.R., Domanski, T.J. and Schloegel, E.L.: "Analysis of Biological Specimens for Volatile Compounds by Gas Chromatography", Journal of Forensic Sciences, Vol. 9, No. 1, January, 1964.

16. Kalant, H. and Kalant, O.J.: "Drugs, Society and Personal Choice", Addiction Research Foundation of

Ontario, General Publishing Co. Ltd., Don Mills, 1972.

17. Firth, P.A.: "Psychologic Factors Influencing the Relationship Between Cardiac Arrhythmia and Mental Load", Ergonomics, 1973, 16 (1): pp 5-16.

18. Sayers, B.: "Analysis of Heart Rate Variability", Ergonomics, 1973, 16 (1): pp 17-32.

19. Ettesna, J.H. and Zielhuis, R.L.: "Physiological Parameters of Mental Load", Ergonomics, 1971, 14 (1): pp 137-144.

20. Opmeer, C.H.J.M.: "The Information Content of Successive RR-Interval Times in the ECG. Preliminary Results Using Factor Analysis and Frequency Analysis", Ergonomics, 1973, 16 (1): pp 105-112.

21. Rohmert, W.: "Heart Rate Variability and Work Load Measurement", Ergonomics, 1973, 16 (1): pp 33-34.

22. Hyndman, B.W. and Gregory, J.R.: "Spectral Analysis of Sinus Arrhythmia During Mental Loading", Ergonomics, 1975, 18 (3): pp 225-270.

23. Kalsbeek, J.W.H.: "Do You Believe in Sinus Arrhythmia", Ergonomics, 1973, 16 (1): pp 99-104.

24. The Technical Cooperation Program, Subcommittee on Non-Atomic Research and Development: "Human Performance and Military Capability in Continuous Operations", December, 1974.

25. Brictson, C.A., McHugh, W. and Naitoh, P.: "Prediction of Pilot Performance: Biochemical and Sleep Mood Correlated under High Work Load Conditions", AGARD Conference Proceedings, No. 146, 1974.

26. Hale, H.B., Hartman, B.O., Harris, D.A., Miranda, R.E. and Williams, E.W.: "Physiologic Cost of Prolonged Double-Crew Flights in C-5 Aircraft", Aerospace Medicine, 1973, 44 (9): pp 999-1008.

27. Blalock, Jr., H.M.: "Social Statistics", McGraw-Hill Book Co., 1972.

28. Philips Electronics Inc. Ltd.: "P-63, Statistical Calculator: Operations Guide", Office Produce Division.

29. Gaito, J.: "Introduction to Analysis of Variance Procedures", MSS Information Corporation, New York, N.Y., 10021.

30. Johnson, L.C. and Naitoh, P.: "The Operational Consequences of Sleep Deprivation and Sleep Deficit", AGARD Graph No. 193: 30, 1974.

31. Goldberg, D.P.: "The Detection of Psychiatric Illness by Questionnaire", Oxford University Press, 1972.

32. Grinker, R.R. and Spiegel, J.P.: "Men Under Stress", McGraw-Hill Book Co., Inc.

33. Frankenhaeuser, M. and Johansson, G.: "Task Demand as Reflected in Catecholamine Excretion and Heart Rate", Journal of Human Stress, Vol. 2, No. 1, March 1972.

34. Gal, R. and Lazarus, R.S.: "The Role of Activity in Anticipating and Confronting Stressful Situations" Journal of Human Stress, Vol. 1, No. 4, December 1975.

35. Laszlo, E.: "Introduction to System Philosophy: Toward a New Paradigm of Contemporary Thought", Harper & Row, 1973.

36. Dexter, L.E.: "Elite and Specialized Interviewing", Northwestern University Press, 1970.

37. Ellis, D.O. and Ludwig, F.J.: "System Philosophy", Prentice-Hall, 1962.

38. Siegel, A.I. and Wolf, J.J.: "Man-Machine Simulation Models: Psychosocial and Performance Interaction", John Wiley & Sons, 1969.

39. Levi, L.: "Society, Stress and Disease: The Psychosocial Environment and Psychosomatic Diseases", Oxford University Press, New York, 1971.

40. Runkel, P.J. and McGrath, J.E.: "Research on Human Behaviour: A Systematic Guide to Method", Holt, Rinehart and Winston, Inc., 1972.

41. Campbell, D.T. and Stanley, J.C.: "Experimental and Quasi-Experimental Design for Research", Rand McNally & Co.

42. Sleight, R.B. and Cook, K.G.: "Problems in Occupational Safety and Health: A Critical Review of Select Worker Physical and Psychological Factors", U.S. Department of Health, Education and Welfare.

43. Selekman, B.M.: "A Moral Philosophy for Management", McGraw-Hill.

44. Churchman, C.W.: "Challenge to Reason", McGraw-Hill.

45. Meehl, P.E.: "Clinical vs. Statistical Prediction: A Theoretical Analysis and a Review of the Evidence", University of Minnesota Press.

46. Rhinelander, P.H.: "Is Man Comprehensible to Man?", Stanford Alumni Association.

47. McLean, A.: "Occupational Stress", Charles C. Thomas Pubs.

48. Appley, M.H. and Trumbull, R.: "Psychological Stress: Issues in Research", Appleton-Century-Croft.

49. Hopkin, V.D.: "Changing Pattern of Stress", CATCA Journal, Fall 1974.

50. Morin, J.D.: "Air Traffic Control: A Demanding Profession Still Without Status", The Controller (IFATCA).

51. Yu, J.C.: "Quantification of Air Traffic Controller's Acceptable Workload", The Controller (IFATCA).

52. Melton, C.E., et al.: "Physiological, Biochemical, and Psychological Responses in Air Traffic Control Personnel: Comparison of The 5-Day and 2-2-1 Shift Rotation Patterns", Federal Aviation Administration, December 1973.

53. Melton, C.E., et al.: "Physiological Responses in Air Traffic Control Personnel: Houston Intercontinental Tower", Federal Aviation Administration, December 1973.

54. Smith, R.C.: "Job Attitudes of Air Traffic Controllers: A Comparison of Three Air Traffic Control Specialties", Federal Aviation Administration, January 1973.

55. Smith, R.C.: "A Realistic View of People in Air Traffic Control", Federal Aviation Administration, December 1974.

56. Wagstaff, A.E.: "Clinical Problems and Stress in Air Traffic Control", The Controller (IFATCA), May 1974.

57. Calhoun, J.R.: "Finding the Right Button", The Journal of Air Traffic Control, July-September 1975.

PHYSIOLOGICAL MEASURES OF WORKLOAD - CORRELATIONS BETWEEN PHYSIOLOGICAL PARAMETERS AND OPERATIONAL PERFORMANCE

by

HELMUT STRASSER

Dr.-Ing.

Institut für Arbeitsphysiologie der Technischen Universität
(Director: o.Prof.Dr.med.W.MÜLLER-LIMMROTH)
Barbarastr.16/I, 8000 München 40, GFR

SUMMARY

Stress defined as input load of man at work may be assessed by means of operational measures and time studies, but strain as the individual output cannot essentially be quantified without also considering physiological data. Variability of some physiological parameters, mostly used in the field of workload studies, will be described and their efficacy discussed for practical applications in field studies. Methodical problems and improvements to counterbalance these problems will be shown, too. Problems involved in interpretations of operational and physiological data, will be dealt with.

Very often in experimental research on stress and strain, correlations between operational, physiological and subjective rating parameters of workload are expected. By means of data from experimental laboratory studies with simultaneously registered physiological and operational performance measures, it will be demonstrated when correlations can be found and when not.

The influence of different hypoxic gas mixtures on pursuit tracking and on some physiological parameters has been studied. From the results the following conclusions can be made: Already in relatively mild hypoxia physiological changes are present, but normally were concealed by reactions due to prolonged test time. Inspite of statistically significant physiological effects, no noticeable deteriorations of performance in tracking could be measured in hypoxia down to a hypoxic gas mixture of only 13% O_2 in inspired air. Not until before 11% O_2 significant and mentionable impairments of tracking performance were found.

The same, shown for hypoxia, is true of noise. Decreased performance in noise could not be found, but an increased level of heart rate indicated the stress.

Hence physiological indicators definitely react already in a low workload in order to bring in action reserves of energy, which will guarantee a normal performance. Often operational measures alone would fail to indicate the strain of the human operator. Only in high workload, correlations between performance and physiological measures can be expected.

INTRODUCTION

In ergonomics research enormous efforts have been undertaken in order to clear to what extent physiological parameters may be used for the assessment of workload. Whereas for scoring physical workload a lot of knowledge and methods have been elaborated, which can be applied already by practitioners, there is a lack of both knowledge and methods in the field of non-physical workload. Inspite of different scientific approaches in physiology and psychology during the last years (see e.g. ROHMERT,1971,1973; KALSBEEK,1967,1971, 1973; VOGT et al.,1973; SCHMIDTKE,1973) a generally adopted method to score objectively influences of mental factors in workload could not be explored. Partly this is due to the variety of different mental load conditions and the fact that human beings, who are stressed by information handling and psycho-social influences, may respond very differently. In this way individual capacities and abilities, by which incoming stresses are coped with, modulate strain variables. But from this relationship between stress and strain the necessity to measure physiological variables when trying to assess strain becomes evident: Normally it is impossible by means of operational data to define to what extent a person is loaded, especially when operational performance still ranges within tolerable limits.

PHYSIOLOGICAL MEASURES OF WORKLOAD

BEAT-TO-BEAT HEART RATE AND SINUS ARRHYTHMIA

As far as continuous physiological measurements are concerned, heart rate as an integral physiological criterion for dynamic and static work, for heat and psychological or emotional stress, is used mostly. But in order to distinguish between the sources of heart rate changes it is necessary to register other parameters, too. E.g. by means of O_2- intake, metabolism of catecholamines or physical measures like temperature or static forces it is possible to find a qualitative analysis of the stress involved. Heart rate is somewhat like an integral over several factors, which determine physiological homeostasis. As an indicator of the "milieu interne" it can be registered in a relatively simple way. Therefore it seemed to be advantageous to deal also with heart rate measurements when trying to assess factors like mental load and fatigue. But heart rate is not an especially suitable parameter for the quantification of mental load in monotonous working conditions which do not require any noticeable motor activities from the subject. Values measured during resting and working phases do not differ systematically (see MULDER & MULDER-HAJONIDES VAN DER MEULEN,1972; STRASSER,1974a).

Because metabolism is not essentially increased in situations which demand concentration and mental activities, heart rate does not react. In addition changes of heart rate resulting from moments of psychological stress mostly are of a very short duration. Therefore they cannot be detected when heart rate is averaged over one minute or even longer time intervals. When heart rate does not change at all under mental load, by means of frequent, dynamic changes in heart rate control, which may be found in the micropicture of heart rate, i.e. in the cardiotachogram, workload may be indicated.

Therefore in order to assess load of the human operator in mainly mental tasks and non-physical workload situations, manifold different scoring methods of the so-called "sinus arrhythmia" (e.g. shown by STRASSER,1974a) have become increasingly important.

Some years ago KALSBEEK (1973) raised the question to himself "Do you believe in sinus arrhythmia?" Of course, this question did not only result from eventually possible scepticism in the parameter sinus arrhythmia but also from discussions with practical men. In the first enthusiasm perhaps they thought, that they had been given by sinus arrhythmia a new commonly valid and powerful measure for assessing mental strain. Yet they might have overestimated the power of prediction of this parameter. If someone is occupied intensively with this question, he may come to the conclusion that sinus arrhythmia as a scoring method should have been a valid one for rather undefined and diffuse situations in workload, i.e. for the quantification of all factors in mental load. In real job designs seldom tasks with an exactly defined mental load are inherent. Additionally disturbing effects of changes in ambient temperature, in different body posture (HANSON & JONES,1970) or in energetic work mostly are involved. Beyond it the relevant information flow, i.e. the load by the job, which is to be quantified as the main cause of the individual's strain, is hidden by a more or less high load of situational information. KALSBEEK (1973) therefore came to the conclusion that sinus arrhythmia at least is "an indicator of the proportional occupation of an individuals single channel capacity during rest and work".

Some examples of the behaviour of heart rate and dynamic changes in heart rate under physical and non-physical stress are shown in Figure 1 and 2. In Figure 1 two recordings are presented which show that increased heart rate and simultaneously decreased variability are the results of increasing energetic work in the course of test time (from EHRENSTEIN,1973). Figure 2 besides other registrations shows beat-to-beat heart rate during rest and tracking as a non-physical workload (STRASSER,1974b). The amplitude of changes in heart rate (arrhythmia) decreases during work while heart rate does even show a slightly lower level during the load of the tracking task.

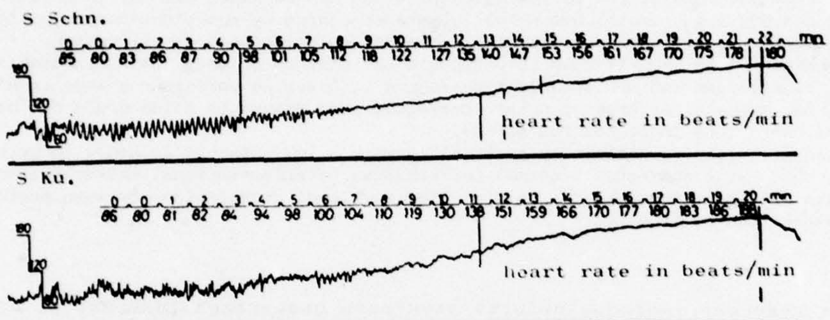


Fig.1: Beat-to-beat heart rate of two subjects (S) influenced by a continuously increasing physical load by means of a bicycle ergometer.

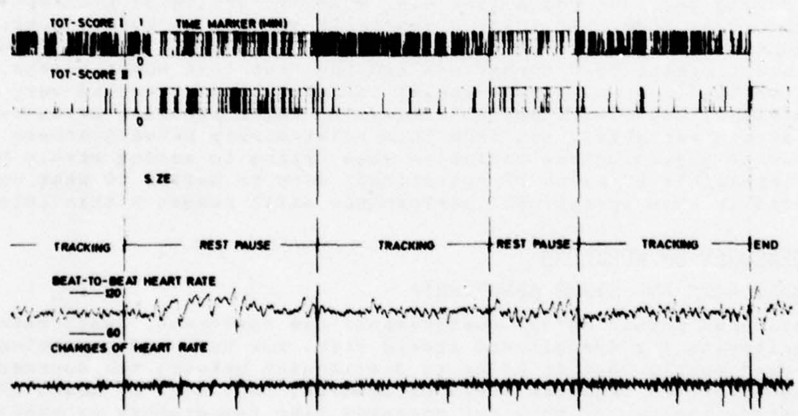


Fig.2: Registration of a cardiotachogram (beat-to-beat heart rate) in rest (rest pauses and end of test) and during a tracking task as a non-physical workload. Besides the cardiotachogram the changes of heart rate from each single beat to the next (changes of heart rate) and different Time-Off-Target scores are shown. By means of the TOT scores performance was measured.

Nowadays it is no problem to measure heart rate and to calculate sinus arrhythmia scores (e.g. sum of absolute differences between succeeding R-top-intervals) by tape recorders or on the basis of telemetering in field studies during a whole shift and to analyze the registrations with a computer. Yet, such profiles should be established by registering heart rate, sinus arrhythmia and e.g. EMG (Electromyogram) of active muscular groups (or EOG (Electrooculogram) in monitoring tasks) not from subjects one after the other, but simultaneously from a group of workers on the same job. By this procedure the disturbing effect of fluctuating situational information flow will be reduced and representations of specific working places in physiological profiles of man can be obtained (compare EINARS, SCHAFFLER & STRASSER,1975; STRASSER & EINARS,1977).

OTHER PHYSIOLOGICAL VARIABLES OF WORKLOAD

Of course other physiological variables which can indicate mental stress reactions (BERKHOUT,1970; LAURIG & ROHMERT,1974) as e.g. EMG or GSR (Galvanic Skin Response) can also be applied in field studies without difficulties, but partly it is not proved whether they are not too complex with regard to interpreting. Some variables, which are proved in laboratory studies, are not for use in field studies because of technical difficulties. E.g. the EEG (Electroencephalogram) may be a spectacular parameter to show slight changes in the activity of the central nervous system. Or by means of evoked potentials, i.e. stimulus-related specific changes in the EEG, it is possible to demonstrate different stages of vigilance (FRUHSTORFER & BERGSTRÖM,1969) and clear effects of alcohol or of relevant and irrelevant information feedback on proficiency and mental load in tracking tests (KLINGER & STRASSER,1972; KLINGER,1973). There has been done very much research in the field of evoked potentials, yet this very delicate parameter is not a practical one for use in field studies. The workers on their jobs would be disturbed too much when evoked responses were elicited. Though flicker fusion frequency (see bibliography from GINSBURG,1970; SCHMIDTKE,1973) is a relatively good indicator of alterations in the vegetative and central nervous system, it can hardly be used while on a job. Determination of the level of catecholamines in the urine (KLIMMER, AULMANN & RUTENFRANZ, 1972,1973) may be a help e.g. in order to decide whether an increased heart rate is due to physical or emotional stress (ULMER,1973), but it must be reminded that such measurements cannot be recorded continuously.

As far as the cardiovascular parameters heart rate and sinus arrhythmia are concerned the following points should be stressed. Heart rate and irregularity scores can be measured continuously as more or less accepted indicators of mental load but still there remains the problem to determine, what changes in physiological parameters do really mean. Are heart rate accelerations of about 10 beats/min during a working phase already dangerous or not? Do increasing or decreasing heart rates signify specific stress which will result in performance decrements? Are there correlations between physiological parameters and performance scores? Has there to be a correlation or not? Because there has not to be a correlation in every case, it is urgently necessary to register physiological parameters, too. This necessity may be elucidated by means of results found during some laboratory studies upon the effects of noise and hypoxia on tracking performance and physiological variables (STRASSER & MÜLLER-LIMMROTH,1973a;1974).

CHANGES IN TRACKING PERFORMANCE AND PHYSIOLOGICAL PARAMETERS UNDER HYPOXIA

METHODS

The influence of hypoxia on tracking and physiological parameters has been researched on a group of 10 young healthy male subjects (*Ss*). Before the tests, Ss have had several trials without hypoxia in order to eliminate effects of learning. During 4 test trials on different days Ss were exposed to defined O_2-N_2 gas mixtures by means of a respiration mask. Each test session consisted of 3 sections lasting for about 45 min, each. The sections included tracking forcing functions of a quasi-fixed and adaptively changing difficulty level (see Figure 3).

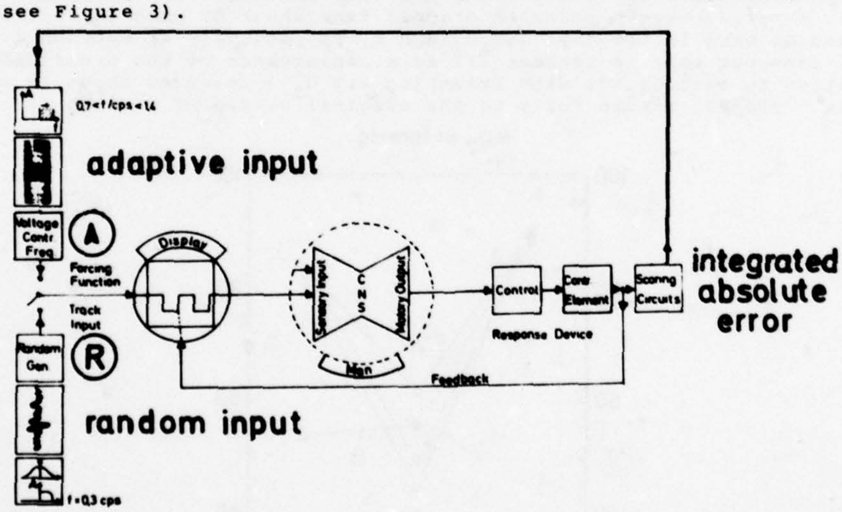
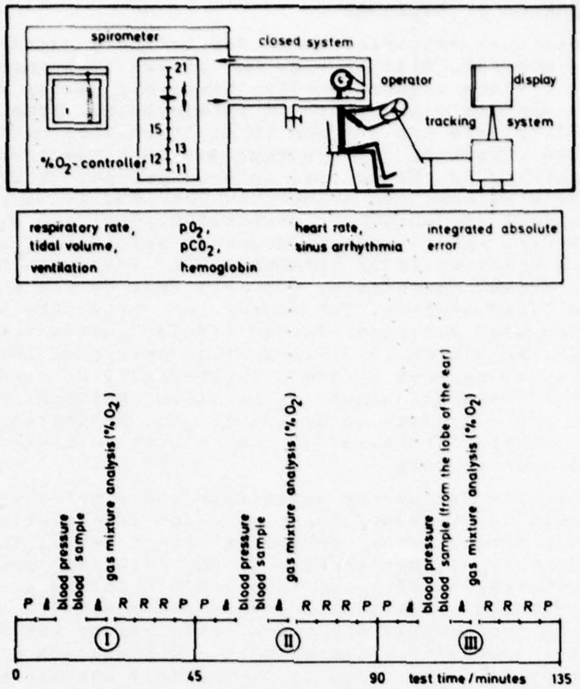


Fig.3: Schematic diagram of the used tracking system with random input signals (Gaussian amplitude distribution with a bandwidth of 0.3 cps; symbolized by R) and adaptive input (progressively increasing frequency of a sinusoidal track, dependent on goodness or poorness of the Ss; symbolized by A).

During resting periods blood pressure was measured and blood samples for analyzing pO_2, pCO_2 and pH, and gas mixtures for control were taken. In a defined work-rest-schedule simultaneously with tracking performance beat-to-beat heart rate, sinus arrhythmia and respiratory parameters were determined (Figure 4).

In a single blind test design Ss inspired normal air (21% O_2) during the first section (I) in each of 4 days. During the second section (II) the O_2-content was reduced to 15%, 13%,12%,and 11% on days A,B,C,and D, respectively. In section III the O_2-content was elevated to 21% on days A,C,and D but held at 13% on day B (compare lower part of Figure 4).



A	21% O_2	15% O_2 – 8,000 ft	21% O_2	A
B	21% O_2	13% O_2 – 12,000 ft	13% O_2	B
C	21% O_2	12% O_2 – 14,000 ft	21% O_2	C
D	21% O_2	11% O_2 – 16,000 ft	21% O_2	D

Fig.4: Outline of the complete experimental design. In a closed-loop spirometry system during 4 different days (A,B,C,D) Ss, who acted as operators in the tracking system, were exposed to normal air in the first (I) and third (III) section and to 4 hypoxic gas mixtures in section II.

RESULTS

Figure 5 presents the mean values of "arterial oxygen partial pressure" under the conditions of normal air (21% O_2 in section I and III) and the 4 different grades of hypoxia in section II. Arterial oxygen pressure dropped from about 90 mmHg in section I to about 60, 55, 50, and 45 mmHg in the days A,B,C, and D, respectively in section II. In day B pressure fell somewhat more in section III as a consequence of the prolonged hypoxia. Oxygen saturation in section III with breathing 21% O_2 - measured about 15 min after the end of hypoxia - did not return fully to the original values of section I.

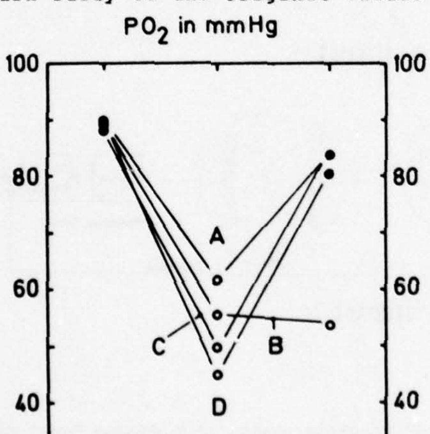


Fig.5: Arterial oxygen partial pressure during the sections I,II, and III of the 4 days A,B,C, and D.

Figure 6 presents tracking performance, which is scored by the integrated absolute error, heart rate and sinus arrhythmia in day A (upper panel left), day B (upper panel right), day C (lower panel left) and day D (lower panel right).

In day A with 15% O_2 during section II no distinct differences between the performance values of the 3 sections can be detected. The same seems to be true of heart rate, but in heart rate an obscured recovery effect can be shown by relating the results of section II and III on section I. Heart rate is not increased from I to II but significantly decreases from II to III.

In day B with 13% O_2 during section II and III absolute error scores are somewhat increased both in section II and III. Heart rate which was depressed significantly from II to III in day A, now is held nearly constant at the same level as in I over the whole test session.

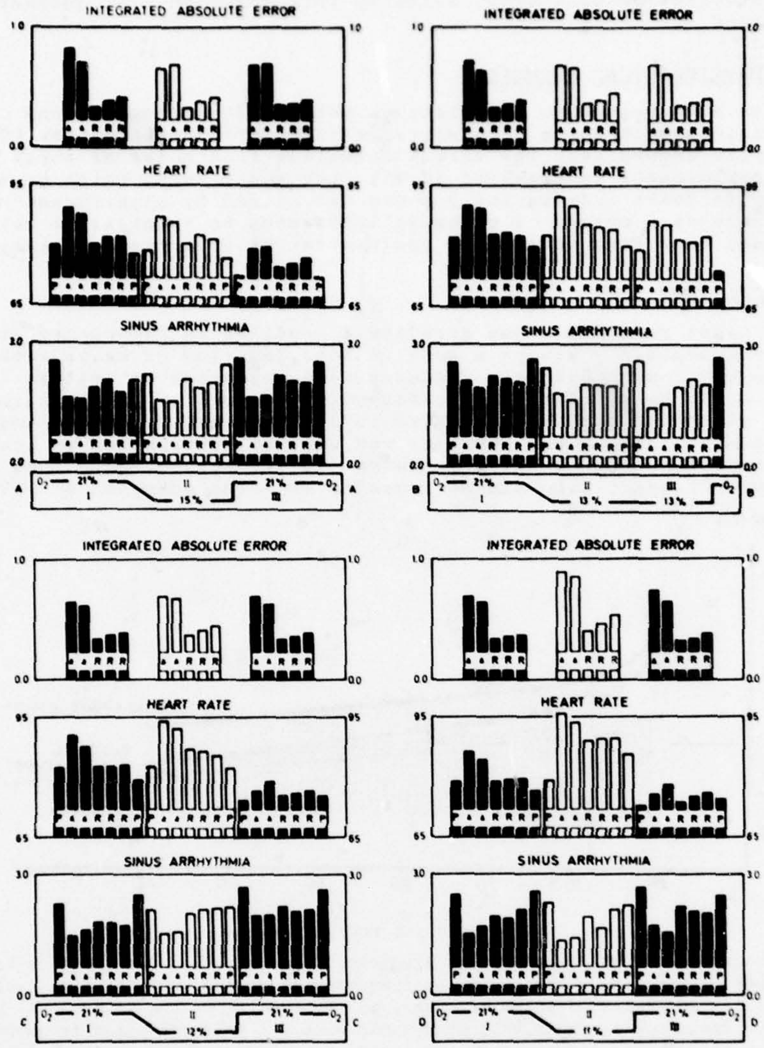


Fig.6: Integrated absolute error, heart rate and sinus arrhythmia during the days A,B,C, and D. The work-rest-schedule is plotted at the bottom of the columns. Rest periods are symbolized by P and the 2 different tracking tasks are marked by **A** for the adaptive test and R for tracking random input signals. Each symbol stands for a duration of 5 min and each column respresents the mean value of 10 Ss. In the lower part of each diagram the level of the content of O_2 in inspired air during the 3 sections I,II and III is plotted. In all diagrams black columns represent periods with normal air.

In day C tracking performance deteriorated as a result of 12% O_2 and heart rate significantly increased in section II and decreased in section III. But the recovery effect, evident in heart rate, is still larger in the change from section II to III than the loading effect of hypoxia from section I to II.

11% O_2 in day D finally result in very strong impairments in all tracking conditions. During tracking random inputs it is obvious that the second and third 5-min error values increase progressively. Also heart rate during all corresponding test phases increases significantly approximately 10 to 15 beats/min.

From the results the following conclusions can be pointed out: With an increasing degree of hypoxia lasting for about 45 min, elevations of heart rate and suppressions of arrhythmia can be expected. Of course arterial oxygen partial pressure as a control parameter of hypoxia is lowered. Already in very mild hypoxia of 15% O_2 hypoxia-related effects in periphery-physiological variables are possible. But if heart rate was only measured under normal air conditions and a section in hypoxia without a following control section, it would be difficult to find the above shown effects. Relaxation in a phase with normal air following the section with hypoxia will give evidence of the stress of preceding hypoxia. Inspite of statistically significant physiological effects no noticeable deterioration of performance could be measured in hypoxia down to 13% O_2. However, with 12% O_2 significant, but still relatively small deteriorations became evident. Only when O_2 in inspired air was reduced to 11% severe impairments of tracking performance of about 30% in the mean became apparent.

Both short time measurements showed a deterioration and with prolonged test time a progressive decline in tracking performance could be seen. Therefore it may be concluded that only high workload results in changes of operational as well as physiological data. Physiological indicators definitely react already in lower workload conditions in order to bring in action reserves of energy. Normally this reaction will guarantee a good operational performance.

NOISE LEVEL AND PHYSIOLOGICAL CHANGES

Similiar results with regard to correlations between physiological and operational variables were yielded with noise as a stressor (STRASSER & MÜLLER-LIMMROTH,1973a). Statistical analysis showed that the stress resulting from noise of about 80 dB(A) did not affect test performance in tracking at all, but heart rate, which generally fell significantly within tests lasting for 3 hours was raised by significant noise-reactive elevations. Probably as a result of relaxing influences no significant noise-induced elevations of heart rate were seen after administration of a tranquilizer.

EFFECT OF PROLONGED TEST TIME

When measuring heart rate and sinus arrhythmia profiles in laboratory studies - as shown in the above chapters - always a more or less lowering of heart rate with increasing test duration was found. But this tendency to degradation as well as the level of sinus arrhythmia could be modulated by different factors. E.g. psychopharmaca which, when stimulating (STRASSER & MÜLLER-LIMMROTH,1973b) hindered declining and when tranquilizing enhanced declining. Such profiles can well be interpreted in accordance with the hypothesis of RUTENFRANZ,ROHMERT & ISKANDER (1971). According to this hypothesis (illustrated in Figure 7) heart rate can be regarded as being composed of several factors.

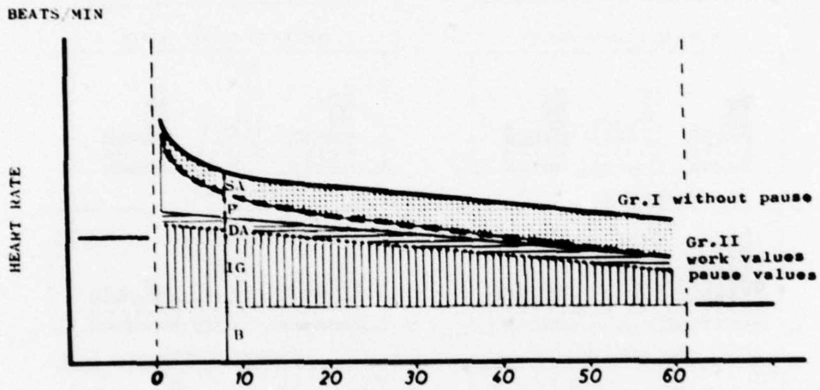


Fig.7: Schematic diagram of the changes of pulse frequency "fractions" on practicing a tracking task after different kinds of practice. B: Basic pulse frequency at repose while sitting (working position) after the trials. IG: Raise of pulse frequency as the result of the intentional basic tension. DA: Raise of pulse frequency as the result of dynamic work. P: Raise of pulse frequency as a result of psychogenic, mostly emotional reactions. SA: Raise of pulse frequency as the result of static work.

Fractions which are called "static" and "dynamic", "intensional basic tension", "basic resting level" and "psychogenic factors" can be distinquished. When applying this hypothesis to the reported results, it seems that the fraction "intentional basic tension" diminished with the length of the sessions. Therefore an overall decreasing of heart rate resulted. This effect of prolonged test time on physiological variables must be taken into account when interpreting experimental results.

CONCLUSIONS

The results with hypoxia and noise , also as possible factors of workload in aerospace medicine, might be very interesting in respect to stress research. No doubt, there had been a very strong stress in the reported tests but no reaction in an operational change occurred. Also in industrial processes very often situations can be met, which are not striking by decrements of performance output of a man-machine-system but by stressed and displeased workers. It stands to reason that in many working situations it will be necessary to show the factors of stress by means of physiological indicators. By this means humanizing of working conditions may become possible.

Another question that arises in this respect is the question of interpreting the results. Concerning this problem it must be said that physiological reactions in field studies have to be calibrated by laboratory experiments. All the more we are not in the fortunate situation to say this or that degree of change is dangerous or will be unhealthy. According to VOGT et al. (1973) a calibration and partitioning of observed heart rates into components, e.g. motor and thermal factors, is possible. By means of standard laboratory tests, motor and thermal cardiac reactivity indices can be obtained. When analyzing heart rates, measured during actual working situations, total reactions can be splitted into the components by the reactivity indices.

REFERENCES

BERKHOUT,J.: Psychophysiological stress: Environmental factors leading to degraded performance. In: Systems Psychology. 407-450. DeGREENE (Ed.) McGraw-Hill Book Company, New York,1970

EHRENSTEIN,W.: Interpretation von Herzfrequenzmessungen mit Hilfe bekannter Beziehungen zu anderen Kreislaufgrößen. In: Pulsfrequenz und Arbeitsuntersuchungen. Schriftenreihe "Arbeitswissenschaft und Praxis". 66-73. Beuth-Vertrieb GmbH., Berlin-Köln-Frankfurt,1973

EINARS,W., SCHAFFLER,K. and H.STRASSER: A computerized automatically system for assessing profiles of preponderant mental load in work study investigations. Pflügers Arch. Europ. J. Physiol.359 (Suppl.)R141(1975)

FRUHSTORFER,H. and R.M.BERGSTRÖM: Human vigilance and auditory evoked responses. Electroenceph. clin. Neurophysiol. 27,346-355(1969)

GINSBURG,N.: Flicker fusion bibliography, 1953-1968. Perceptual and Motor Skills 30,427-482(1970)

HANSON,J.A. and F.P.JONES: Heart rate and small postural changes. Ergonomics 13(4)483-487(1970)

KALSBEEK,J.W.H.: Mentale Belasting. Theoretische en experimentele exploraties ter outwikkeling van meetmethoden. Assen, Van Gorcum, TNO,1967

KALSBEEK,J.W.H.: Sinus arrhythmia and the dual task method in measuring mental load. In: Measurement of man at work. 101-113. SINGLETON,FOX & WHITFIELD (Eds.)Taylor & Francis, London,1971

KALSBEEK,J.W.H.: Do you believe in sinus arrhythmia? Ergonomics 16(1)99-104(1973)

KLINGER,K.-P.: Amplitudenvariationen akustisch evozierter Potentiale in Abhängigkeit von der Signalinformation und belastungsbedingter Ermüdung. Int. Z. angew. Physiol. 31,269-278(1973)

KLINGER,K.-P. and H.STRASSER: Variations of physiological parameters during defined mental load and rest. Pflügers Arch. Europ. J. Physiol.332(Suppl.)R82(1972)

KLIMMER,F., AULMANN,H.M. und J.RUTENFRANZ: Katecholaminausscheidung im Urin bei emotional und mental belastenden Tätigkeiten im Flugverkehrskontrolldienst. Int. Arch. Arbeitsmed. 30,65-80(1972)

KLIMMER,F., AULMANN,H.M. und J.RUTENFRANZ: Katecholaminausscheidung bei mental belastenden Tätigkeiten im Flugverkehrskontrolldienst. In: Problematik von Arbeitsplätzen mit mentaler Belastung. Pathogene Stäube mit ihren Auswirkungen auf den Menschen. 69-75. WENZEL & TENTRUP (Hrsg.). Bericht über die 12. Jahrestagung der Deutschen Gesellschaft für Arbeitsmedizin,1972, Gentner Verlag, Stuttgart,1973

LAURIG,W. und W.ROHMERT: Ergonomische Methoden zur Beurteilung des Teilsystems "Mensch" in Arbeitssystemen. In: Ergonomie 2. 113-145. SCHMIDTKE (Hrsg.) Carl Hanser Verlag, München,1974

MULDER,G. and W.R.E.H.MULDER-HAJONIDES VAN DER MEULEN: Heart rate variability in a binary choice reaction task: An evaluation of some scoring methods. Acta psychologica 36,239-251(1972)

ROHMERT,W.: Introduction to: An *International Symposium on* Objective Assessment of Work Load in Air Traffic Control Tasks. Held at the Institute of Arbeitswissenschaft, The University of Technology, Darmstadt, GFR. Ergonomics 14(5)545-547(1971)

ROHMERT,W.: Pulsfrequenz und Dauerleistungsgrenze. In: Pulsfrequenz und Arbeitsuntersuchungen. Schriftenreihe "Arbeitswissenschaft und Praxis". 21-33. Beuth-Vertrieb GmbH., Berlin- Köln- Frankfurt, 1973

RUTENFRANZ,J., ROHMERT,W. und A.ISKANDER: Über das Verhalten der Pulsfrequenz während des Erlernens sensumotorischer Fertigkeiten unter besonderer Berücksichtigung der Pausenwirkung. Int. Z. angew. Physiol. 29,101-118(1971)

SCHMIDTKE,H.: Mentale Beanspruchung. In: Ergonomie 1. 256-279. SCHMIDTKE (Hrsg.) Carl Hanser Verlag, München,1973

STRASSER,H. und W.MÜLLER-LIMMROTH: Komplexe Auswirkungen der Faktoren Lärm, Tranquilizer, erschwerte Arbeitsbedingung und Versuchszeit auf eine Pursuit Trackingleistung und das kontinuierliche Puls-zu-Puls-Verhalten. Int. Arch. Arbeitsmed. 31,81-103(1973a)

STRASSER,H. und W.MÜLLER-LIMMROTH: Physiologische Veränderungen und Regelleistungsverhalten älterer Probanden während kontinuierlicher Trackingtätigkeiten nach Zufuhr einer zentral aktivierenden Substanz. Arzneim.-Forsch. (Drug.-Res.) 23(3)406-415(1973b)

STRASSER,H.: Beurteilung ergonomischer Fragestellungen mit Herzfrequenz und Sinusarrhythmie (Indicatoren von mentaler Beanspruchung und Ermüdung). Int. Arch. Arbeitsmed. 32,261-287(1974a)

STRASSER,H.: Technisch-physiologische Aspekte der Beziehung Stress - Strain. Eine modell-theoretische Betrachtung. Arbeitsmed.-Sozialmed.-Präventivmed. 9(10)212-217(1974b)

STRASSER,H. and W.MÜLLER-LIMMROTH: Tracking performance and physiological parameters in hypoxia. 22. Int. Congress of Aviation and Space Med. Abstract of papers,p12,Beirut/Lebanon,6.-12.10.,1974

STRASSER,H.: Physiological measures of mental load. In: Symposium of Production and Methods of Stress Research Association (PROMSTRA),Kalmar/Sweden, 25.-28.9.,1974. Taylor & Francis, London,1977

STRASSER,H. und W.EINARS: Beanspruchungsprofile in der ergonomischen Feldforschung. Arbeitsmed.-Sozialmed.-Präventivmed. 12(1)6-10(1977)

ULMER,H.V.: Physiologische Grundlagen zur Beurteilung der Arbeitsbeanspruchung mit Hilfe von Pulsfrequenzmessungen. In: Pulsfrequenz und Arbeitsuntersuchungen. Schriftenreihe "Arbeitswissenschaft und Praxis". 41-50. Beuth-Vertrieb GmbH., Berlin-Köln-Frankfurt,1973

VOGT,J.J., MEYER-SCHWERTZ,M.Th., METZ,B. and R.FOEHR: Motor, thermal and sensory factors in heart rate variation: A methodology for indirect estimation of intermittent muscular work and environmental heat loads. Ergonomics 16(1)45-60(1973)

DISCUSSION

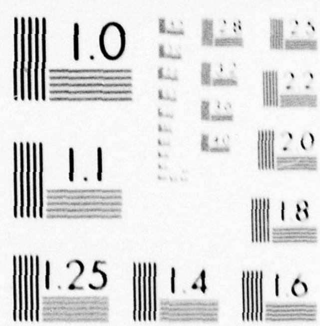
HARTMAN: (United States)	I just want to comment on your use of sleep logs. We, too, employ this sort of instrument and find that it is easy to use and generates quite useful data. You have employed it in a most effective fashion.
BRICTSON: (United States)	Do you have objective data on fragmented sleep?
SOUTENDAM: (Canada)	No. This is subjective, but we have found it to be quite adequate.
GREEN: (United Kingdom)	Do you have any personality data on those people taking drugs, etc.?
SOUTENDAM: (Canada)	No. We haven't obtained that sort of data.
SANDERS: (Netherlands)	Some of this drug use may be due to the environmental stress and individual sensitivity to stress of air traffic controllers.
SOUTENDAM: (Canada)	I agree. However, I won't go so far as to say controllers take drugs because they are under stress.
OFFENLOCH: (Germany)	You have doubts about biochemical measures of stress. Can paper-and-pencil tests differentiate between various kinds of stress?
SOUTENDAM: (Canada)	No. These tests simply help you decide whether you need to go to more sophisticated measures.
NICHOLSON: (United Kingdom)	Did the control group also report disturbed sleep? I think this is rather important.
SOUTENDAM: (Canada)	Yes, they did.

AD-A057 835 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 5/9
METHODS TO ASSESS WORKLOAD.(U)
JUN 78

UNCLASSIFIED AGARD-CP-216 NL

2 OF 2
AD
A057 835

END
DATE
FILMED
10-78
DDC



MICROCOPY RESOLUTION TEST CHART

USE OF INSPIRATORY MINUTE VOLUMES IN EVALUATION OF ROTARY AND FIXED WING PILOT WORKLOAD

Frank S. Pettyjohn, LTC, MC, USA
Roderick J. McNeil, CPT, MSC, USA
Lloyd A. Akers, SP6, Card. Pul. Tech, USA
James M. Faber, SP4, Bio. Tech, USA
US ARMY AEROMEDICAL RESEARCH LABORATORY
FORT RUCKER, ALABAMA 36362
USA

SUMMARY

Inspiratory Minute Volume (IMV) measurements by Mueller Respirometer were utilized in the evaluation of US Army aircrew workload and stress in helicopter and fixed wing aircraft. The IMV data obtained demonstrates a significant stress and/or workload level of the aviator in performance of helicopter day nap-of-the-earth (NOE), night nap-of-the-earth flight (NNOE) and with the use of night vision devices (NVD). IMV of 20.05 to 38.11 liters per minute NTPD were obtained during the performance of these combat operational techniques. IMV determination in-flight is considered a valuable clinical tool in the assessment of aircrew stress and/or workload.

BACKGROUND

The operational employment of US Army helicopters and fixed wing aircraft in the high intensity air defense environment has required the development of advanced flying techniques. Day nap-of-the-earth (NOE), night nap-of-the-earth (NNOE), and night vision devices (NVD) have added markedly to the workload of the aircrew.

During oxygen utilization studies, analysis of Inspiratory Minute Volume (IMV) data indicated significant trends dependent on flight profiles. To further evaluate this method as an indirect and simple modality in the assessment of stress and workload, a study of IMV during varied helicopter and fixed wing aircraft flight profiles was undertaken. The term IMV is utilized for simplicity in place of the usual pulmonary function notation $\dot{V}_I$.

MATERIALS AND METHODS

Data was obtained from a series of 135 flights of OH-58, UH-1H, and AH-1 helicopters and U-21 and C-7 fixed wing aircraft. Four phases of the flight were evaluated--runup (R.U.), takeoff (T.O.), cruise (C), and final approach (F.A.). The helicopter flight profile was evaluated under normal flight, NOE, NNOE, and NVD. Thirty (30) subjects wore standard US Army A-13A oxygen masks connected to a portable Mueller respirometer. Inspiratory minute volumes obtained in liters per minute (LPM) were corrected to NTPD which is normal temperature (70°F), pressure (760 Torr), and dry. Barometric pressure was obtained by on-board barometer accurate to 0.5 Torr. The volume of inspired air was measured throughout the period of time in each phase of flight. The longer phases are cruise and during NOE, NNOE, and NVD. The minimum period for any sample was 10 minutes. Temperature of inspired air was determined by standard thermometer intrinisic to the Mueller respirometer.

RESULTS

Fixed Wing Data (U-21, C-2). The IMV data from evaluation of fixed wing flight profile are provided in Table I.

TABLE I

FIXED WING INSPIRATORY MINUTE VOLUME (NTPD)

		U-21				C-7			
Flight Condition		R.U.	T.O.	C.	F.A.	R.U.	T.O.	C.	F.A.
NORMAL FLIGHT	Mean	6.82	10.21	8.19	14.19	3.33	9.76	8.25	19.12
	S.D.	2.44	0.93	3.36	9.30	---	---	0.32	---
	Range	5.10-8.55	9.65-10.77	3.63-12.79	3.11-34.06	---	---	8.06-8.51	---

Table I demonstrates the IMV response to the flight profile. Of note is the increased IMV at takeoff and final approach for fixed wing aviators.

Rotary Wing Data (OH-58, UH-1H, AH-1). The IMV data from flights of three helicopter types under four mission profiles is shown in Table II.

The rotary wing IMV demonstrates the aviator's perception of his aircraft and mission profile threat. The highest threats perceived are the nap-of-the-earth profiles and the night vision devices.

DISCUSSION

The US Army aviator's perception of his aircraft is reflected in his IMV. The phase of flight evaluated produces a wide range of response from the baseline IMV. Fixed wing aircraft (U-21 and C-7) studied during routine flight operations indicate takeoff and final approach for landing are the periods of greatest stress or workload. This finding is in consonance with that of the USAF and USN.[1] The IMV obtained exceeded the current US military design standard of 13.12 LPM NTPD for oxygen systems during the final approach.[2]

During rotary wing flight, the two primary variables affecting IMV are the type helicopter and mission profile. The AH-1G Cobra gunship is perceived as a "constant" threat under the NOE flight profile. The IMV remains an approximate constant of 15 LPM NTPD throughout all phases of flight.

The OH-58 and UH-1H during a routine flight profile demonstrate an essentially stable IMV at expected levels of 9 to 11 LPM NTPD. The Army aviator exposed to the threat and workload of NOE, NNOE or NVD profiles demonstrates significant increase in his IMV. During NOE and NNOE in the UH-1H, the IMV is increased twofold. The use of the NVD in the UH-1H provides extreme values for IMV throughout the flight profiles. A threefold increase (35 LPM NTPD) occurs in cruise and final approach phase while using the NVD.

Interestingly, the final approach phase of day NOE in the UH-1H is viewed as a release from threat or stress as evidenced by a decrease in IMV to less than baseline levels. With use of the NVD, final approach is perceived as an extreme stress as indicated by the threefold increase. This finding supports the use of actual in-flight IMV determinations to differentiate aircrew workload from workload and stress/threat conditions.

The values of IMV obtained under the increased stress of NOE, NNOE and NVD far exceed the current oxygen design standards. Comparison of the mean values of 20 to 38 LPM NTPD (22.5 to 42.6 LPM BTPS) obtained in this study to the USN standard of 23.7 LPM BTPS or the USAF value of 25.1 LPM BTPS demonstrates the limitations of present US military oxygen design standards.[3,4]

The IMV values provide the initial estimate of the level of aircrew stress involved in current helicopter operations.

CONCLUSIONS

Stress of helicopter operations under the advanced operational concepts of nap-of-the-earth, night nap-of-the-earth flight, and the use of night vision devices has been evaluated by the use of inspiratory minute volume determinations. The IMV obtained during NOE, NNOE and NVD is increased threefold demonstrating the aviator's perception of the increased stress and/or threat. The IMV data obtained in the routine flight of US Army fixed wing aircraft is consistent with the USAF and USN data documenting the stress and increased workload during takeoff and landing. Use of IMV is considered a valuable tool in the clinical assessment of aircrew stress and workload.

REFERENCES

1. Morgan, T. R., Reid, D. H. and Baumgardner, F. W. Pulmonary Ventilation Requirements Evident in the Operation of Representative High Performance Aircraft. Aerospace Medical Association Scientific Meeting Preprint, 156-159, 1976.

2. Military Specification--Design and Installation of Gaseous Oxygen Systems in Aircraft, General Specification for. MIL-D-8683A, 3 July 1969.

3. Jackson, M. M. Usage Rates of Aircraft Oxygen. WADD Technical Report 60-106, 1960, Wright-Patterson AFB, Ohio.

4. Roman, J., Olden, H. and Jones, W. L. Flight Research Program VII: Medical Monitoring of Navy Carrier Pilots in Combat. Aerospace Medicine 38:133-137, 1967.

TABLE II

ROTARY WING INSPIRATORY MINUTE VOLUMES (LPM, NTPD)

		UH-1H					OH-58				
Flight Condition		R.U.	T.O.	C	T	F.A.	R.U.	T.O.	C	T	F.A.
ROUTINE	Mean	---	11.59	10.95	---	10.72	9.67	9.06	9.19	---	9.32
	S.D.	---	0.75	0.47	---	3.24	3.12	2.95	3.03	---	1.80
	Range	---	11.06-12.12	10.48-11.42	---	7.04-13.14	6.30-19.67	7.31-18.96	5.13-16.22	---	7.56-12.83
							AH-1G				
NOE	Mean	---	12.61	16.23	20.050	7.71	---	15.97	15.17	14.49	15.73
	S.D.	---	10.12	2.69	0.98	1.87	---	3.39	5.38	4.72	2.06
	Range	---	1.11-25.37	14.81-22.22	17.56-21.41	8.15-10.45	---	12.23-18.84	11.37-18.98	10.89-19.14	13.34-17.08
NNOE	Mean	---	---	11.88	21.45	17.20	---	---	---	---	---
	S.D.	---	---	2.13	7.054	1.75	---	---	---	---	---
	Range	---	---	9.44-14.16	13.15-32.66	15.96-18.94	---	---	---	---	---
NVD	Mean	---	18.20	---	34.51	38.---	---	---	---	---	---
	S.D.	---	0.33	---	0.71	0.94	---	---	---	---	---
	Range	---	17.91-18.67	---	53.69-34.92	37.13-39.01	---	---	---	---	---

NOTE: All values Liters Per Minute, Normal Temperature 70°F, Pressure 760 mmHg, Dry (LPM, NTPD); S.D. = Standard Deviation; R.U. = Runup; T.O. = Takeoff; C = Cruise; T = Threat (NOE, NNOE, NVD); F.A. = Final Approach; NOE = Nap of Earth; NNOE = Night Nap of Earth; NVD = Night Vision Device.

NEUROPHYSIOLOGICAL ASSESSMENT OF FUNCTIONAL STATES OF THE BRAIN

by

Kurt Offenloch
Neurophysiologisches Laboratorium
Zentrum der Physiologie
Klinikum der J.W. Goethe – Universität
Theodor Stern – Kai 7
6000 Frankfurt am Main, FRG

ABSTRACT

By neurophysiological methods functional states of the brain can be assessed with a precision that surpasses that of classical psychological methods.

Not only neurological syndromes, but also slight changes on the vigilance scale as well as functional changes associated with cognitive and intellectual functions can be correlated with the electrical activity of the brain and hereby objectively determined.

INTRODUCTION

The examination of functional states of the brain under normal and pathological conditions involves techniques and procedures of neurosurgery, neurology, neurophysiology, neurochemistry, psychiatry and psychology. Since neurosurgery, neurology and psychiatry are mainly dealing with diseases of the brain and mind, the investigation of the working brain under physiological conditions is mainly performed by neurophysiology, neurochemistry and psychology.

This paper deals with one of the most important neurophysiological methods: electroencephalography (EEG), which together with electromyography (EMG) is forming also the basic methods of clinical neurophysiology. These methods occupy a unique position in brain research since they constitute an objective approach and are noninvasive. Measurements can be repeated several times without distress for the patient or subject. They are indispensable to neurologists and neurosurgeons in the diagnosis of many neurological diseases.

Brain function is associated with the production of electrical potentials. The recording of this electrical activity of the brain from the scalp and its study is called electroencephalography. It was discovered by the German psychiatrist Hans Berger in 1924. Since 1929, the year of Berger's first publication of his discovery (Ref.1), electroencephalography grew into a new and important science since the EEG reflects different functional states and dispositions to reactions of the brain.

Today the classical period of conventional visual evaluation and interpretation of the EEG signals written as curve traces on the paper of an EEG machine has definitely reached its ultimate stage. No further developments of the art and science of EEG interpretation by visual analysis can be expected.

New insights into the basic processes of EEG wave generation and EEG correlates of information processing within the brain have come by the advent of the application of microelectrode, stereotaxically implanted multi-electrodes and computer techniques for recording and analyzing the EEG signals. These methods which were developed in the last two decades in various leading laboratories in Europe and North America gave us extremely powerful tools for solving many before unsolvable aspects of the enigmatic EEG wave processes.

ELECTROGENESIS OF THE EEG

A better understanding of the basic wave processes of the EEG, its electrogenesis, was obtained by neurophysiological qualitative and quantitative analyses of the relationship between EEG and single nerve cell activity within the cerebral cortex. Initiated by the introduction of intracellular micro-electrode recording techniques in brain and cerebral cortex physiology, a close relationship between the time course of EEG potentials and intracellular slow membrane transients of cortical neurons, particularly membrane potential changes due to depolarizing and hyperpolarizing synaptic

potentials have been found by several investigators (Refs. 2,9). Their investigations resulted in the following modern concept of EEG electrogenesis: the EEG potentials result from the temporal and spatial integration of the extracellular cerebral potential fields of spontaneous and evoked membrane potential fluctuations of neuronal – and possibly to an unknown extent also of glial – elements. A further outcome was that only the activity of larger ensembles of synchronized neural elements is detectable in scalp recordings.

COMPUTERIZED EEG ANALYSES

New and most important impulses were given to the quantitative EEG analysis by different techniques of computation derived from signal analysis (e.g. averaging of stimulus-dependent evoked responses (Ref.11), correlation and spectral analysis) which in the last three decades were introduced parallel with the fast development of electronic data processing and computer technology also in the analysis of the EEG. They allowed a much more detailed examination of the electroencephalogram and resulted in a definite gain of information as compared with the conventional visual analysis of the EEG signal. Recently two symposia dealt with different methods and applications of computerized EEG analyses (Refs. 5, 8).

APPLICATIONS

The well-known clinical applications of electroenchephalography, or the EEG correlates of metabolic states and metabolic functions of the brain as revealed under the pathophysiological conditions of comatose states, exogenic and endogenic intoxications or hypoxia will not be covered by this presentation. Also the assessment of sleep stages or of cerebral age for example will not be discussed here.

I want to restrict myself to applications which – in my opinion – are more relevant to the topics of this meeting.

1. Consciousness and Vigilance

Consciousness is bound to intact functions of the cerebral cortex and subcortical structures (e.g. Formatio reticularis and the so-called unspecific thalamic nuclei), the latter which are physiologically activating the cerebral cortex. So disturbances of consciousness not only occur by cortical lesions but also as a result of lesions or functional disturbances of lower brain areas (brain stem, thalamus, limbic system).

Already Berger found that the EEG changes with altered consciousness. With some exceptions one generally can state that the slower the EEG, the more serious the disturbance of consciousness. Changes in the level of vigilance are shown in the EEG prior to clinical signs, and small decreases in vigilance which can not be detected by psychological test scales can be detected by evaluation of the EEG records (Ref.4).

2. Genetically Determined Different Learning Performance

EEG patterns are strongly genetically determined. Also the most intensively studied electrophysiological correlate of learning has been especially the EEG activity of the hippocampus, a subcortical structure which according to experimental and clinical reports is considered as essential in the learning process and the fixation of the mnemonic trace.

The EEG of the sensorimotor cortex and of the dorsal hippocampus was recorded during basal conditions before, and three, five and ten days after the beginning of training in two strains of rats with genetically determined different conditioned behaviour in an active two-way avoidance test. One group of animals were "good" learners, the other group were "bad" learners.

Spectral analysis with a frequency resolution of 0.5 Hz was performed on EEG epochs of 16 sec duration. From 20 consecutive single spectra also an average spectrum was computed. Of the 20 spectra of 16 sec duration (that is 320 sec) an intercorrelation matrix of 2 Hz segments was computed. In that intercorrelation matrix certain frequency bands emerge since adjacent spectral segments – if belonging to the same frequency band – are significantly correlated with each other. The complete intercorrelation matrix was submitted to a factor analysis by which the most important spectral segments were thus obtained.

This statistical evaluation of the spectral components of the EEG displayed differences in the distribution of the frequency bands and their relative intensities in "good" and "bad" learners before and after training. So it was possible to distinguish different EEG patterns resulting from genetic and conditioned influences. (Ref.6).

3. Cognitive and Intellectual Functions

Cognitive and intellectual or mental functions are bound to the cerebral cortex. It is well established that changes of EEG activity take place during mental performance in man. Generally, a "synchronized" EEG is recorded in a relaxed state with eyes closed, whereas a "desynchronized" EEG is recorded when a subject becomes mentally active.

Lehmann (personal communication) found that different EEG characteristics are associated with the execution of different tasks: arithmetic, reading, cerebral compared with spatial mental tasks, eye/target vs hand/target coordination.

According to Giannitrapani (Ref.7) power spectra of the EEG differentiated not only between conditions differing in the sensory modalities but also between auditory tasks involving patternless sound (white noise), patterned nonverbal sound (music), patterned verbal sound (listening to a story). All conditions were clearly and significantly different in several cortical areas and frequency bands. By the same author also correlations between EEG spectra and the scores of Wechsler Intelligence Scale were obtained for 11 to 13 year old children with clusters of significant correlations in different cortical areas and for different frequency bands.

In my laboratory Waldeier (Ref.12) dealt with the electroencephalographic differentiation of cognitive functions of man such as "reading" and "mental arithmetic" in comparison with the resting conditions with open and closed eyes. We were especially interested in the changes of spectral intensities within the four EEG frequency bands (alpha, beta, theta and delta) under the four investigated psychophysiological states which could be clearly distinguished by computerized EEG analysis.

4. Psychophysiological Load

Quite recently it was shown by Walz (Ref.13) that the compterized EEG also can be used in a man-machine control system for the assessment of the degree of psychophysiological load of the controller man at the execution of control tasks.

In a driving simulator three control tasks had to be performed, one with a low, the other with an average, and the third with a high degree of difficulty. A clear relationship between the control performance, the subjective estimation of the degree of difficulty of the control task, and the EEG spectral values could be demonstrated.

REFERENCES

1. Berger, H. — *Über das Elektrenkephalogramm des Menschen.* I. Mitteilung. Arch. Psychiat. Nervenkr., 87: pp.527-570, 1929.

2. Creutzfeldt, O.D. — *Neurophysiological Correlates of Different Functional States of the Brain* in: Brain Work, Alfred Benzon Symposium VIII, Munksgaard, pp.21-47, 1975.

3. Callaway, E. — *Brain Electrical Potentials and Individual Psychological Differences.* Grune Stratton, New York, 1975.

4. Heimann, H., Schmocker, A.-M. — *Zur Korrelation frequenzanalytischer und psychologischer Messwert* in: Quantitative Analysis of the EEG. Methods and Applications. Proceedings 2nd Symposium Study Group EEG-Methodology, pp.631-647, Jongny sur Vevey, 1975.

5. Dolce, G., Künkel, H., Eds. — *CEAN-Computerized EEG Analysis.* Symposium Merck'sche Gesselschaft für Kunst und Wissenschaft, Kronberg/Taunus, 1974. Gustav Fischer, Stuttgart, 1975.

6. Dolce, G., Offenloch, K., Sannita, W.G., Müller-Calgan, H., Decker, H. — *EEG-Untersuchung des Trainingseffektes bei gut und schlecht lernendent Ratten.* Statistische Analysen der Spektralwerte. Arch. Psychiat. Nervenkr., 217: pp.139-148, 1973.

7. Giannitrapani, D. — *Spectral Analysis of the EEG* in: CEAN – Computerized EEG Analysis, pp.384-402, Gustav Fischer, Stuttgart, 1975.

8. Matejcek, M., Schenk, G.K., Eds. — *Quantitative Analysis of the EEG. Methods and Applications.* Proceedings 2nd Symposium Study Group EEG – Methodology, Jogny sur Vevey, 1975.

9. Offenloch, K. — *What is the EEG Composed of? Aspects of EEG Electrogenesis* in: CEAN – Computerized EEG Analysis, pp.71-84, Gustav Fischer, Stuttgart, 1975.

10. Offenloch, K., Dolce, G. — *Brain Functions Investigated by Computerized EEG Analyses.* Proceedings First International Congress Biological Psychiatry, Buenos Aires, 24–28 September 1974.

11. Regan, D. — *Evoked Potentials in Psychology, Sensory Physiology and Clinical Medicine.* Chapman and Hall, London, 1972.

12. Waldeier, H. *Elektroenzephalographische Differenzierung kognitiver Tätigkeiten bei gesunden Personen.* Ein Beitrag zur Anwendung der Spektralanalyse des EEG und weiterer statistischer Analysen der EEG – Spektralwerte. Dissertation, Frankfurt am Main, 1974.

13. Walz, L. *Dynamisches Regelverhalten und hirnelektrische Vigilanzregulierung des Menschen bei der Durchführung von Regelaufgaben.* Dissertation, Technische Universität Berlin, 1976.

DISCUSSION

MONESI: (Italy) There are rather sophisticated EEG analyses which have been developed to determine the level of psychotropic drive.

OFFENLOCH: (Germany) Yes, I agree, and I feel that the analytic techniques used here fit into that category.

MONESI: (Italy) You could apply these kinds of measures to the selection problem if you had baselines.

OFFENLOCH: (Germany) Yes; if there are baselines available. However, the method I have just reported can be applied much more widely.

CANNINGS: (United Kingdom) Can you describe the statistical techniques used to establish the significant differences you have reported here?

OFFENLOCH: (Germany) No, I was not the statistician in this project. I can't give you any more information than that which I have just reported.

NICHOLSON: (United Kingdom) You have shown some rather specific EEG changes, but your conclusions are more far-reaching. For example, you state that the EEG can show effects better than psychological tests. I see no evidence for your statement.

OFFENLOCH: (Germany) I will simply point out that the EEG can monitor man continuously and will reveal slight changes. Psychological tests can't do either of these things.

CLEMENT: (Belgium) Your method of analysis is basically fast Fourier. What is the minimum length of signal analyzed?

OFFENLOCH: (Germany) The analysis segments are 2, 4, 8, or 12 seconds. For the most part, we used 12 seconds.

CLEMENT: (Belgium) Did you also take into account coherence measures?

OFFENLOCH: (Germany) We have done coherence studies but haven't found anything in that approach which is useful to us.

CLEMENT: (Belgium) Did you complete your spectral analysis assessment by means of pattern recognition?

OFFENLOCH: (Germany) I have no pattern recognition data, but believe there are special types of pattern recognition.

GRATZL: (Germany) Did you use EEG to look at responses to different color patterns?

OFFENLOCH: (Germany) No, we have not done that.

SANDERS: (Netherlands) Have you reconfirmed the arithmetic results with variations in motivation?

OFFENLACH: (Germany) There is no EEG pattern for motivation.

SANDERS: (Netherlands) Does your EEG analysis indicate the state of arousal or the kind of task?

OFFENLACH: (Germany) The EEG reveals the cognitive state of the subject.

THE HUMAN OPERATOR SIMULATOR: WORKLOAD ESTIMATION USING A SIMULATED SECONDARY TASK

by

Norman E. Lane, Ph.D.
Lieutenant Commander, Medical Service Corps, U.S. Navy
Naval Air Development Center
Warminster, Pennsylvania 18974

and

Melvin I. Strieb
Robert J. Wherry, Jr., Ph.D.
Analytics, Inc.
2500 Maryland Road
Willow Grove, Pennsylvania 19090

During the design phase of system development, it is necessary to examine emerging system configurations to minimize potential operator overload problems. Two general classes of workload assessment procedures have been applied toward this goal. The first of these approaches, analytic prediction or time/motion models, is intended for use early in the design cycle and involves the estimation, for each task, of demands on discrete processing channels and of estimated time to perform specific task components. These models are useful in identifying gross overload conditions early enough for corrective action to be initiated. A second class of workload assessment approaches is designed for application when the system has progressed to the dynamic simulation stage. While data yielded by these simulation assessment procedures is more precise than that supplied by techniques used in early design, development and application costs are high, flexibility is limited, and output is usually too late in the development cycle to allow the correction of system deficiencies.

This paper describes the use of the Human Operator Simulator (HOS) for identifying potential workload problems. HOS is applicable during the midrange of system development, between early analytic prediction methods and later simulator evaluations. The HOS system creates a software simulation of a trained human operator, his system hardware/software and required interfaces. Previous HOS applications have indicated a close correspondence in the behavior of HOS operators and human operators on the same tasks. The results of this work demonstrate the applicability of HOS for workload evaluation and explores HOS operator behavior under varying conditions of task demand. Problems of definition and methodology for current workload measures are discussed and alternatives identified using HOS to control task difficulty and task demand parameters. The flexibility of HOS allows differing system designs to be evaluated with neither hardware modification nor operator retraining. The implications of HOS application for cost and flexibility improvements are examined, and further developments of the model for workload evaluation are proposed.

INTRODUCTION

Modern aircraft systems and their increasingly sophisticated avionics have imposed extremely complex performance requirements on system operators. Demands on the operator routinely approach and often exceed the upper limits of operator capability. Within some highly automated crewstations, crewmembers are able to exercise only a fraction of the repertoire of functions provided by onboard computers; at the same time they are often overwhelmed by having to perform numerous other functions that are frequently more adaptable to automation than those already computerized. Problems in operator loading have become an area of deep concern for human factors in the development of new systems; unfortunately, this area remains one of the most resistant to satisfactory solution.

Difficulties in the resolution of operator loading problems commence during early design phases and continue throughout the system life cycle. These difficulties stem in large part from the rapidity with which critical design decisions are made during very early stages of development. Long before mockup or simulator validation of crewstation design can occur, the majority of decisions which affect the crewstation have become essentially irreversible, regardless of the outcome of such validation. Systems become operational with established crew overload problems. Subsequent versions of the system address the problem with more and better avionics and automation, but seldom consider the basic question of the operator and his capabilities as part of an integrated system. To interface with such new "labor saving" avionics capability, crewmembers inherit increased mission responsibility, and the overload cycle continues.

These rapidly increasing operator demands have affected the system roles of all classes of operators. Lovesey (1) reports a twofold increase in the number of fighter aircraft instruments between 1960 and 1970. Given the trend within this time frame toward multiple-purpose displays, it is probable that the actual number of items of information presented to the operator increased at least 300 percent during that period, and is still rising. It is most common in discussing workload to treat pilot overload as the key to systems performance. For many modern airborne tactical systems, however, the pilot has only moderate impact on mission success and may be the most lightly loaded crewmember. The traditional focus on piloting behavior in workload studies has ignored the critical problems that exist for acoustic and non-acoustic sensor operators, radar operators and tactical decision makers. For these crewmembers, information and activity overload is chronic and generally unrecognized. Procedures which are developed to predict or measure workload in emerging systems must be sufficiently flexible to accommodate a broad variety of system tasks and further, must be applicable during early design before key task allocations are fixed.

Workload Measurement

In response to well-established difficulties with attaining acceptable crew loading in operational systems, considerable research has been directed by human factors engineers, ergonomists, physicians, physiologists, psychologists, and control theorists toward the goal of defining and measuring the elusive quantity of "operator loading." Schiflett (2), in a highly selective bibliography, cites more than 80 references dealing directly with the operator workload concept. Jahns (3,4) analyzes and categorizes numerous workload measurement efforts. Rolfe (5) provides a systematic discussion of the advantages and disadvantages of different loading measures. These discussions point out the dramatic diversity of approaches and definitions employed by workload investigators. The term "workload measures" must encompass such dissimilar concepts as spectral analysis of control inputs, cross-adaptive secondary tasks, and complex physiological measures of activation and arousal.

Workload assessment approaches differ in key aspects, such as the characteristics of the principal measurement variables. Howitt (6) addresses the dimension of the time frame over which operator behavior is observed; the period of interest may range from a few seconds during a specific mission phase, to a week or more, incorporating fatigue and circadian rhythms into the workload question. Another distinction among these approaches, perhaps the most critical for the systems designer, is the point at which they can be applied during the development cycle. Applicability of workload measurement efforts spans the whole system life cycle, from estimates of workload based on preliminary function allocation to quantification of operator output during actual system operations. In general, as system development proceeds, task requirements become more clearly defined, and measurement of workload becomes more precise. At the same time, the utility of study findings diminishes until little or no impact on system design is possible. Rolfe (5) stresses the need for multiple approaches throughout the development cycle, with relatively imprecise measures required in early development and more refined assessments applied as simulators or operational prototypes become available. This need for a broad spectrum of techniques to address measurement of operator effort is supported by Jahns (3), who posits four categories for current workload research and cites illustrative studies. The class of approaches called "Time/Motion Studies" has traditionally been of greatest concern and of greatest value in early design. These models, sometimes called analytic prediction methods, are based on task analyses and task timelines derived from early function allocations. The prototype of such models was developed and refined by Siegel and Wolf (7,8). A similar approach was developed by Dickey (9) and refined into the Workload Assessment Model (WAM) by Boeing under U.S. Navy sponsorship as part of the Computer Aided Function Allocation-Evaluation System (CAFES) (10). A description and application of the WAM methodology is presented by Linton, Jahns and Chatelier (11) elsewhere in these proceedings. Another outgrowth of the Siegel-Wolf approach with application to operator loading is the Systems Analysis of Integrated Network Tasks (SAINT) developed by the U.S. Air Force (12). Still another in this class of analytic approaches in early development is Time-Based Load Analysis (13,14), which employs mission time-line analyses to estimate operator load. Wingert (15) suggests a modification to such analysis-based approaches which accommodates the apparent ability of skilled operators to perform certain types of tasks simultaneously. In this function-interlace approach, the interference between processing channels which inhibits task simultaneity is quantified through a matrix of interlace coefficients. All these approaches and their modifications can be considered as task-analytic methods of predicting or estimating workload; all require the estimation of numerous parameters describing task performance times and demands imposed on specific processing channels by each task. To the extent that these estimates are satisfactory, analytic prediction or time/motion studies can be of considerable utility in detecting gross design deficiencies and potential overload conditions sufficiently early for design changes to be implemented. They are, however, broadly approximative methods and are appropriate primarily in the context of early design.

After the system has progressed to the simulation/prototype stage, another broad class of approaches becomes applicable. Although measurement philosophy and procedures vary considerably within these approaches, all share certain common characteristics. All require the construction of a functioning hardware/software simulator or prototype, and all require a human operator trained in the tasks to be performed on the simulator. Numerous measurement paradigms have been applied in these simulator-based studies. Stackhouse (16) and Spyker and colleagues (17) describe an approach using combinations of physiological measures and an adaptive performance task. Johannsen, Pfendler and Stein (18) demonstrate the use of a combination of multiple measures of system performance, ratings and a secondary task. Smit (19) posits a similar approach to an inflight measurement problem. These examples illustrate the broad and relatively vague use of the term "workload" even when applied within a relatively narrow portion of the development cycle.

A further problem with simulator-based approaches is the necessity for the involvement of a human operator. Flexibility and adaptability to changing conditions, and intuitive approaches to problem-solving make the human a superior component of a man-machine system. These same characteristics make the human a less-than-ideal subject for the investigator examining a set of system procedures for potential overload problems. A given operator may or may not be thoroughly trained. He may be more or less inherently capable than another operator. His motivation to perform is uncontrolled. His definition of satisfactory performance on each of the many component tasks will be unlike that of any other operator. His adaptability will lead him to perform the tasks however he can, taking shortcuts and changing procedures to cope with task demands, leading sometimes to incorrect conclusions that system procedures are adequate when, in fact, they are only adequate if a skilled, motivated operator ignores them. The lack of a "standard" operator can cause a massive lack of sensitivity to critical design deficiencies during simulator evaluation of design.

Analytic prediction methods function as highly sophisticated deterministic calculators, summarizing and reorganizing the task data provided them. They are inexpensive to run and can be repeated for major system modifications to evaluate highly divergent alternative designs. Simulator-based methods are applicable late in system development. They provide highly realistic simulations of the behavior of hardware and software components, but are dependent on the uncertainties of human performance to achieve satisfactory system validation. They are relatively expensive to develop and operate, and major system modifications or repeated runs on alternate designs are usually impractical. There are thus substantial gaps between the two approaches, both in philosophy and in the point of applicability to system development. Between

early design assessment by analytic models and late system validation by simulators, there is only one currently available technique which attempts to evaluate the effects of task characteristics and crewstation design on operator performance. The HOS model has the capability of simulating, by computer, the functions of the hardware/software components of a system and the behavior of a human operator working within the proposed system. HOS is described in more detail in the following sections.

THE HUMAN OPERATOR SIMULATOR

HOS is a technique for crewstation design and evaluation which simulates the behavior of a trained human operator performing a specified mission within a defined crewstation layout. In support of the HOS operator, the model also allows simulation of the functioning of system equipment and the required interfaces to the operator. The rationale and development history of the HOS system is reported by Wherry (20). Among the unique features of the HOS approach is the use of a number of "micromodels," each of which controls the simulated execution of a specific operator function such as reaching, control actuation, encoding/decoding, absorbing and recalling information, and decision making. Each micromodel effects changes in system status or in an operator state, and assesses "time charges" against the mission in accordance with its own set of rules derived from human performance data and specific experimentation. These basic operator functions can be combined to simulate virtually any discrete or continuous task; tasks can be combined into sets of tasks called procedures. This process is performed by another part of the HOS system, HOPROC, a computer programming language with internal algorithms for aggregating micromodels, allowing the writing of operator tasks and procedures in a near-English language format. Procedures are tied together through a series of multiplexed control routines and through a "banker" which collects time charges and records system transactions for later analysis by HODAC, an analysis program which collects, analyzes, and reports simulation outcomes in a variety of formats.

The same procedures language (HOPROC) used to code operator behavior can be used to code simulations of system components. Given a system input/output specification, HOS can closely reproduce the transfer of information between operator and system and the operations performed on that information by any system component. HOS is thus capable of simulating performance in highly complex tasks, and allows precise control of the task priorities used by the simulated operator with respect to order of execution during multitask demand situations. HOS has successfully simulated a number of relatively simple tasks, such as reach performance, multiple dial monitoring and mail sorting (21) and a complex operational mission, that of the Air Tactical Officer in the LAMPS antisubmarine helicopter (22). An additional feature of HOS, only partially implemented, is the specification of operator states (O-States) which allow the user to vary a number of basic operator performance capabilities. O-States provide for examination of individual differences in operators or for the simulation of degraded environmental conditions by causing functions to be performed with an increase or decrease in operator efficiency.

HOS and Workload Assessment

One of the principal categories of workload research outlined by Jahns (3) is the class called "information processing" studies. Work in this area is characterized by the use of secondary tasks to assess additional operator capability beyond that required by a primary task. The performance attained on the secondary task is taken as an indication of reserve capacity available; the higher this performance without primary task decrement, the lower the primary task workload is presumed to be.

Despite the conceptual attractiveness of the secondary task or "divided-attention" approach, numerous problems have emerged in its application to workload assessment. These problems are, in general, the same as those discussed above with respect to simulator workload evaluation. Human operators differ in capability and task familiarity. Their motivation is unknown. Most critically, differences in interpretation of instructions can result in the use of widely different strategies and produce uncontrolled inter-operator differences in length and duration of transfers to the secondary task. A clear technical presentation of these methodological difficulties is given by Kahnemann (23).

It is proposed that the use of a simulated operator can alleviate a significant portion of these methodological problems. A principal virtue of the HOS operator is a willingness to do exactly as "he" is instructed. The capability within HOS of establishing numerous task parameters allows a wide range of operator behavior to be elicited without the uncontrolled variation produced by divergent operator strategies. Task difficulty can be accurately varied, and the length and duration of task interrupts can be precisely manipulated. In addition, the concept of "internal limits" employed by the model allows the operator's definition of "satisfactory performance" to be specified as a parameter. This control insures that the primary task performance will not vary as a result of decreased operator effort or shortcuts as secondary task demands increase. More detailed discussion of the behavior of the HOS operator under divided-attention demands and of the control capability provided by model parameters is contained in (24).

Prior to the presentation of the four HOS workload studies, it is necessary to discuss how the term workload might be quantified. Standard practice is to interpret a deterioration in primary task performance when a secondary task is introduced as evidence that the secondary task increased the operator workload. It is quite another thing to use data such as the primary task's relative or absolute performance deterioration and the relative time spent on the secondary task to arrive at a quantitative measure of the workload of the primary task. One type of equation which has been suggested for use in computing workload in divided-attention studies is of the general form

$$\text{Workload} \leq \frac{P'}{P} - T \qquad (1)$$

where P = performance on the primary task when the secondary task is not present,
P' = performance on the primary task when the secondary task is present, and
T = the proportion of time spent on the secondary task.

From Eq. (1), it can be seen that if no degradation occurs on the primary task, then the workload of the primary task cannot be higher than 1-x, where x is the fraction of time spent on the secondary task.

While Eq. (1) may appear to offer a logical and clever approach to quantifying workload, it raises at least three potentially major difficulties:

(1) Does the specific measure of operator performance used in the calculation influence the result of the equation?

(2) Assuming that the percent of time spent on the secondary task could be held constant, would the frequency and duration of transfers to the secondary task (the pattern of accumulated secondary task time) affect performance on the primary task?

(3) Do the task characteristics of the secondary task affect performance on the primary task?

The first point deals with the "external criterion" of success and represents an arbitrary selection by the workload investigator. Only if the results of Eq. (1) are insensitive to this type of decision (which has little, if anything, to do with either actual task difficulty or the level of operator effort being expended) should one posit that the measure is a quantitative measure of the workload.

The second point raises the issue of whether the workload of a given task can be considered a viable concept. If the result of Eq. (1) changes as a function of the pattern of task interrupts, then numerically equal percentages of time spent on the secondary task in two different studies will not be equivalent, and each primary task will have to be represented by a family of workloads rather than a single figure. When such multiple workload curves are required, single workload measures are much less accurate than commonly supposed, and may be misleading if the task conditions selected are not representative of conditions actually encountered by an operator.

The third point raises the issue that Eq. (1) does not consider the type of secondary task being accomplished, but is affected only by the percentage of time spent away from the primary task.

Raising such questions as these about such familiar terms as task demand or operator load or workload, focuses attention on the fact that these terms are merely intervening variables (like "fatigue" or "motivation") which appear to be highly useful ways of talking about the general problem of an operator having "too much to do," but, when examined closely, defy strict operational definition and, thus, encourage over-simplification of a highly complex problem.

The studies to be described were formulated to investigate some of these issues. HOS (with its ability to control task difficulty and eliminate or control individual differences in such factors as operator fatigue, motivation, and interpretation of instruction) can closely replicate task conditions in divided-attention studies involving secondary tasks. It is thus possible to investigate the effect of such parameters as primary task difficulty, percent of time spent on the secondary task, and the frequency and duration of interrupts by the secondary task on primary task performance, as measured by a variety of arbitrary external criteria. Whether or not these studies support conventional approaches to quantifying workload, their outcomes are at least unaffected by uncontrolled variation between subjects or between tasks.

The HOS system was designed as a tool for prediction of operator behavior and analysis of operator performance. Although operator loading measurement is implicit within such a simulation, aspects of the model which are useful for workload evaluation have not been made explicit in previous applications. The intent of the work described in the remainder of this paper has been,

(1) to demonstrate the applicability of HOS to problems of the type encountered in a multi-task workload assessment context,

(2) using a dual-task simulation, to explore the behavior of the HOS operator under varying conditions of task demand and task performance criteria, and

(3) to examine HOS measurements of workload which can be used to compare alternative design configurations.

DIVIDED-ATTENTION STUDIES USING HOS

Approach

The primary task selected for these studies was a standard one-dimensional pursuit tracking task. It was felt that such a continuous type of task would show greater sensitivity to minor degradation in operator performance in high demand situations than would a discrete task.

In the primary task, the operator was presented with a display that indicated two values -- the signal to be tracked and the track. The display was calibrated from 0 to 20,000 units. The track was controlled by a rotary control knob; a one-degree turn of the knob changed the track by 200 units. The signal varied in a pseudo-random fashion, and was computed by summing three sine functions with different amplitudes and frequencies. While a human does learn to anticipate (predict) the behavior of a single sine wave, when several sine waves of different frequencies, amplitudes and phases are combined, the human is unable to learn to anticipate the resulting signal. Thus, such tasks are recognized as being highly similar to real-world tracking problems where the target signal is likely to exhibit highly unpredictable behavior.

The operator was instructed to keep the track within a predefined range around the signal. An interfering secondary task periodically interrupted the primary task and kept the operator from working on the primary task for a fixed period of time. Both the frequency and duration of transfers to the secondary task were controlled.

The secondary task(s) may be thought of as either a continuous or discrete task that, while being worked on, required the operator's complete attention. For a real human operator this period of forced

(in)activity could be effected by blanking out the display; in HOS, the simulated operator was simply given an explicit instruction to wait a specific amount of time before resuming work on the primary task. Thus, the secondary task was totally neutral, avoiding the possibility of any task-specific interactions arising between the primary and secondary tasks. The secondary task can be likened to the interruptions that a combat pilot must schedule to check his angle-of-attack or range-to-target, or to select or arm or launch his weapon while maneuvering the aircraft into a favorable target attack position (the primary task). Each secondary task results in a short interruption to the primary task. The secondary task is considered to be just as critical as the primary task. Therefore, it must be performed at the appropriate time and cannot be delayed.

The Tracking Task Model

A simple tracking task model might assume that the operator repeatedly executes a basic adjustment cycle consisting of:

(1) Reading the value of the signal

(2) Reading the value of the track

(3) (a) Terminating the adjustment cycle if the difference between the two values is less than the defined tolerance limit

(b) Computing the amount of the knob turn required to make the track equal to the signal (using the precise gain factor that relates knob movements to changes in the track signal) and making the appropriate change, if the difference exceeds the tolerance limits.

This simple tracking model, however, is clearly unrealistic in that it:

(1) Assumes the operator knows the precise gain factor of the system

(2) Fails to recognize the highly adaptive character of human tracking performance.

Therefore, in order to make the model more plausible, a more complex model was developed in which the operator:

(1) Learns to extrapolate the value of the signal to compensate for the tendency to lag behind the signal

(2) Learns the gain factor to use in his calculation of knob adjustments so as to minimize the error he perceives in his tracking performance.

These learning processes are expressed as mathematical functions in which the new extrapolation times and gain factors are determined from old values of these variables and from estimated values of the signal and track obtained on the current and previous adjustment cycles.

During the course of model development, a variety of algorithms for both extrapolation time and gain learning were investigated. Many apparently reasonable algorithms, it was found, produce unstable learning, i.e., the learned values tended to diverge rather than converge to an optimal value. The learning algorithms used in the simulation experiments to be described, however, have been found to produce stable learning in all situations of interest. A complete description of these learning algorithms as well as a discussion of the stability problem is presented elsewhere (24). It should also be emphasized that these learning models are intended to represent the continual learning of a practiced subject presented with a signal whose characteristics may be continually changing, rather than the learning process of an operator who is initially unfamiliar with the task and/or the hardware system.

Measures of Operator Performance

Two performance criteria, time-on-target and RMS error, are used in tracking studies. While both are legitimate, the former tends to be more useful in system performance studies and the latter in laboratory performance studies. Time-on-target assumes that a "window" around the signal can be defined such that if the operator has the track within the window, he is on target, and if it is outside the window, he is not. The operational interest is not on how far outside (or inside) the window he is, but on the amount (or percent) of time-on-target. In this series of studies, it was decided to use both types of criteria but to include five different window sizes. The window sizes employed are discussed in Study 1.

Study 1 -- The Effect of Task Difficulty and Time-Sharing on Tracking Performance

Many experiments have indicated that, in general, as the frequencies and amplitudes of the component sine waves that form the signal become faster and larger, respectively, it becomes more difficult for an operator to maintain performance, i.e., operator workload increases, in some sense. Study 1 was designed to examine the effect of variations in the amplitudes and frequencies of the component sine waves on the various performance measures. In addition, it examined the effect that addition of a secondary task had on performance under each of the amplitude-frequency combinations.

The results of Study 1 are shown in Table 1. Three amplitude-frequency combinations shown as Tasks 1, 2, and 3, were used. In Task 1, each of the three component sine waves had amplitudes equal to half those comprising Task 3 (4000, 2800, and 1600 for Task 3). In Task 2, each of the three component sine waves had frequencies that were half the Task 3 values ($2\pi/10$, $2\pi/7$ and $2\pi/4$ for Task 3). Each task was run both without the secondary task and with a secondary task that required 40 percent of the operator's time (six seconds on the primary task followed by four seconds on the secondary task). Time-on-target was

TABLE 1. HOS WORKLOAD ASSESSMENT SUMMARY DATA

STUDY NO.	TASK NO.	% TIME	CYCLE TIME	TIME ON PRIMARY	TIME ON SECONDARY	INTERNAL LIMITS	RMS ERROR	PERCENT TIME ON TARGET 1600	800	500	300	100	CONTROL ADJUSTS
1	1	100.0	6.0	6.0	0	800	455	99.4	93.7	74.7	48.9	17.5	396
	1	60.0	10.0	6.0	4.0	800	1417	81.6	65.1	44.4	28.7	9.7	173
	2	100.0	6.0	6.0	0	800	562	99.9	82.7	56.3	35.8	13.2	215
	2	60.0	10.0	6.0	4.0	800	2625	71.5	59.0	40.2	24.3	8.4	155
	3	100.0	6.0	6.0	0	800	910	91.7	71.5	50.4	32.8	11.3	625
	3	60.0	10.0	6.0	4.0	800	3357	66.5	51.2	37.7	23.9	8.1	309
2	3	60.0	2.5	1.5	1.0	800	3004	52.7	31.0	20.5	12.5	3.7	360
	3	60.0	5.0	3.0	2.0	800	3243	59.9	42.1	27.6	16.3	4.8	301
	3	60.0	7.5	4.5	3.0	800	3613	65.9	44.5	30.3	18.8	6.7	317
	3	60.0	10.0	6.0	4.0	800	3357	66.5	51.2	37.7	27.9	8.1	309
	3	60.0	12.5	7.5	5.0	800	3523	65.7	49.7	33.4	20.7	6.7	293
	3	60.0	15.0	9.0	6.0	800	3713	65.2	49.0	34.3	21.5	7.4	313
3	3	100.0	6.0	6.0	0	800	910	91.7	71.5	50.4	32.8	11.3	625
	3	85.7	7.0	6.0	1.0	800	1582	83.2	64.7	43.8	27.6	9.5	402
	3	75.0	8.0	6.0	2.0	800	2109	77.3	58.3	40.9	25.4	9.0	362
	3	66.7	9.0	6.0	3.0	800	3083	70.4	52.1	36.2	21.5	6.9	321
	3	60.0	10.0	6.0	4.0	800	3326	66.8	51.5	37.9	24.0	8.2	309
	3	54.5	11.0	6.0	5.0	800	3968	59.2	44.8	30.9	18.4	6.0	243
	3	50.0	12.0	6.0	6.0	800	4318	56.9	42.4	30.5	18.5	6.2	220
4	3	100.0	6.0	6.0	0	800	910	91.7	71.5	50.4	32.8	11.3	625
	3	75.0	8.0	6.0	2.0	800	2109	77.3	58.3	40.9	25.4	9.0	362
	3	75.0	8.0	6.0	2.0	600	2041	78.6	62.3	45.6	29.8	10.9	442
	3	75.0	8.0	6.0	2.0	400	2123	78.3	60.2	44.1	30.0	12.1	511
	3	75.0	8.0	6.0	2.0	200	1999	79.7	61.4	47.9	32.0	11.0	591
	3	75.0	8.0	6.0	2.0	0	1923	78.6	62.1	49.6	35.9	12.9	696
	3	60.0	10.0	6.0	4.0	800	3357	66.5	51.2	37.7	23.9	8.1	309
	3	60.0	10.0	6.0	4.0	600	3224	64.2	48.0	35.2	22.4	7.5	343
	3	60.0	10.0	6.0	4.0	400	3228	66.5	51.4	40.0	25.9	9.2	383
	3	60.0	10.0	6.0	4.0	200	3111	66.5	50.3	39.7	27.9	10.3	446
	3	60.0	10.0	6.0	4.0	0	3393	67.4	52.1	42.0	30.9	11.7	569

calculated with respect to five criteria -- the percents of time that the track was within 1600, 800, 500, 300 and 100 units of the input signal. In all studies, each task was performed for 200 seconds of simulated time.

The results show clear performance degradation with respect to all criteria -- both RMS error and the various time-on-target percentages increase as the amplitude and frequency increase and when the secondary task is added. However, the effects on the RMS error and time-on-target have no predictable, functional relationship to the actual changes in amplitude or frequency.

It should be noted that the addition of the secondary task has a clear effect on both RMS error and time-on-target but that merely noting the percent increase in RMS error (or decrease in time-on-target) would not enable the difficulty of the tasks to be ranked meaningfully.

Study 2 -- The Effect of Length of Alternation Periods on Primary Task Performance

With both HOS and real operators, internal strategy algorithms determine when the operator will shift from one task to another. However, HOS gives the investigator the ability to control the exact amount of time the simulated operator can (or must) work on any given task (as well as which task he must work on next). This feature was used in Study 2 to study the effect that variation in the frequency with which the operator alternated between the primary task and the secondary task would have on performance in the primary task.

It is clear that the longer the operator is away from the primary task the less likely it will be that the last track position will be satisfactory and the greater will be the overall error. On the other hand, the longer the operator has to make corrections once he returns to the primary task, the greater the likelihood that he will have time to resume satisfactory tracking performance of the target. Thus, variation in the alternation period could conceivably lead to degraded, similar, or improved performance depending on other characteristics of the task.

Study 2 examined the effect of varying the alternation periods on RMS error and time-on-target. For this study, the amplitude-frequency combination from Study 1, Task 3 was used. A constant ratio of time on primary to time on secondary was maintained; 60 percent of the time was spent on the primary task and 40 percent on the secondary. The primary-secondary alternation cycle was varied from 2.5 seconds to 15 seconds in 2.5-second increments.

The results of this study are shown in Table 1 and plotted in Figure 1. Both RMS error and time-on-target are curvilinear functions of the length of the alternation period. For all five time-on-target criteria, there is an optimum value of the alternation period that represents the best trade-off between the factors mentioned above -- time away versus time available to perform corrective actions for this particular task. At approximately the same time, RMS error seems to level off near its asymptote. However, the alternation period associated with the best performance by the criterion of RMS error is the period associated with the worst performance according to all of the time-on-target measures. This demonstrates that the selection of either RMS error or time-on-target criteria as measures of operator workload must

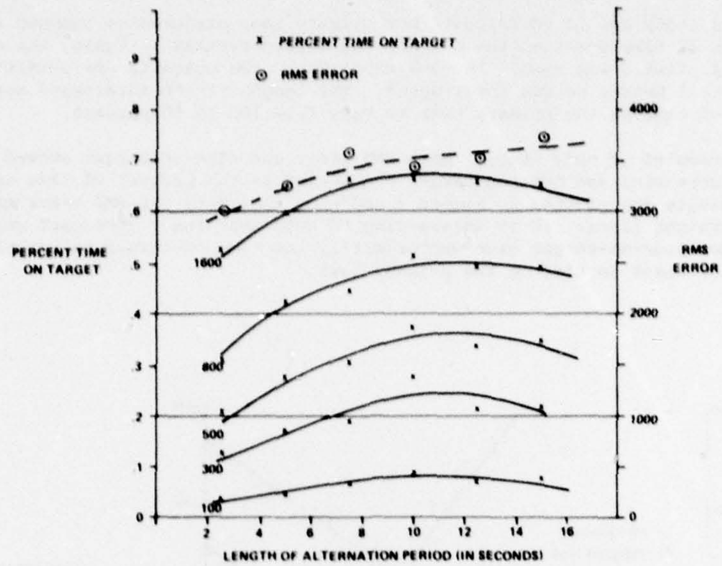


FIGURE 1. ALTERNATION PERIOD VS. RMS ERROR AND TIME ON TARGET

be carefully made on the basis of their appropriateness for a particular situation because they may give contradictory results.

The data was further analyzed to determine the rates of deterioration in time-on-target and recovery as a function of time away from and back on the primary task. It was found that, for all conditions, the deterioration in the values of time-on-target could be enclosed within the envelope shown in Figure 2 and that recovery followed the trajectories shown by the dotted lines.

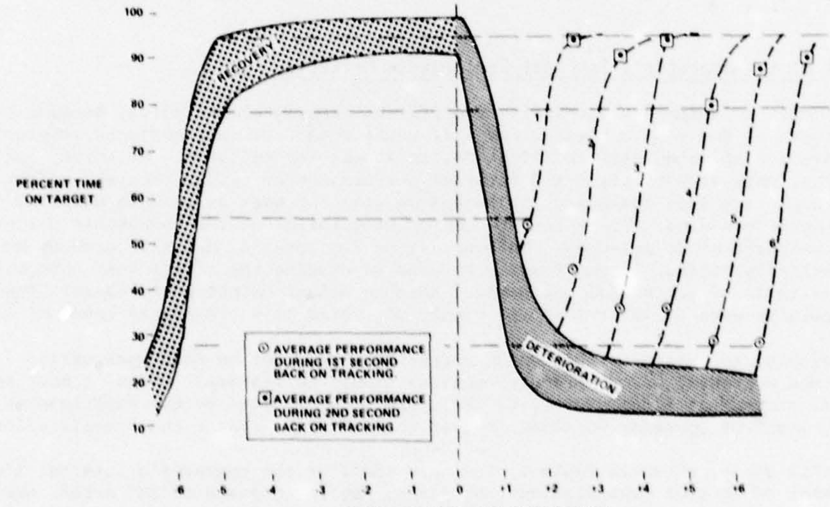


FIGURE 2. SECONDS FROM INTERRUPTION OF TRACKING VS. TIME ON TARGET

Study 3 -- The Effect of Percent Time on Secondary Task on Tracking Performance

The purpose of this study was to investigate how primary task performance changed as a function of the relative percentages of time spent on the primary and secondary tasks. Again, the amplitude-frequency combination from Study 1, Task 3 was used. In each experiment, the operator was permitted to work for six seconds on the primary task before he was interrupted. The length of the interrupts were varied causing the overall percentage of time on the primary task to vary from 100 to 50 percent.

Table 1 shows the results of this study. Both RMS error and time-on-target showed consistent variations, with RMS error increasing and time-on-target decreasing as the percent of time spent on the primary task decreased. The results are plotted in Figure 3 and show that both the RMS error and time-on-target values can be fit by straight lines. It is interesting to note that the y intercept points extrapolated from the fitted lines for time-on-target give approximately the time-on-target values that would be expected by chance if the operator spent no time on the primary task.

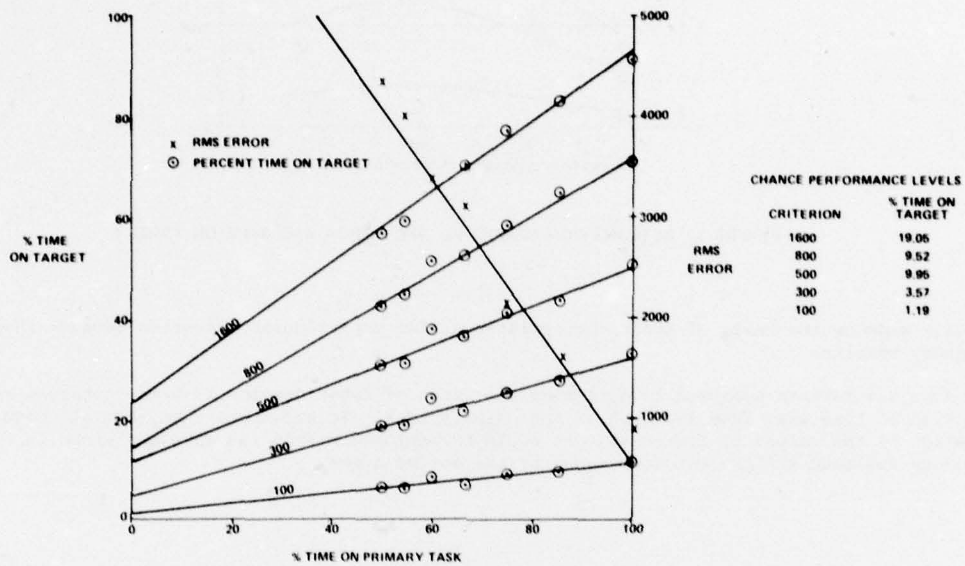


FIGURE 3. PERCENT TIME ON PRIMARY TASK VS. PERCENT TIME ON TARGET AND RMS ERROR

Study 4 -- The Effect of the Operator's Internal Performance Criteria

It is often difficult to compare the results of different experimental studies, because they do not use either the same tasks or the same subjects. Even if studies use the same subjects, the order in which the various studies are accomplished can, and often do, influence the outcomes. Moreover, individual differences among subjects, such as motivation and internal performance criteria, severely affect the results. For example, operators who are less demanding of themselves will not work as hard to minimize the error as operators who are more demanding. These factors can greatly influence the operator's perceived "task demand" and the various performance measures. However, since the internal criteria used by HOS can be easily and deterministically varied, it can readily be used to examine the effect that changes in the operator's internal criteria of successful performance have on actual output performance. Thus, HOS operated as if the operator were fully trained and highly motivated to a predefined level of achievement.

In Study 4 the display tolerances within which corrections need not be made were varied. Clearly, as the tolerances become narrower, more corrective actions should be required. Thus, a more restrictive set of limits, whether arrived at independently by the operator or imposed by the experimenter, should tend to increase the amount of operator workload, but at the same time reduce the overall error.

The results of this study, shown in Table 1, indicate that, as the operator's internal limits criteria are narrowed, the number of control manipulations increases. By the measure of RMS error, there is no significant change, i.e., the operator maintains the same overall performance. By the criterion of time-on-target, however, there is a significant change in performance, with the time-on-target increasing as the limits are narrowed. The effect seen in these runs may be likened to the operator's reserve capacity enabling him to perform the same task more precisely within the same amount of time by working harder.

In this regard, it is likely that an operator may use a different set of internal limits when performing a single continuous tracking task than when there were other tasks to accomplish. For an easy primary task, a relatively wide set of limits may produce acceptable performance when there are no secondary tasks. However, when confronted with one or more secondary tasks, the operator may have to consciously or subsonsciously narrow the limits on the primary task in order to maintain performance. Thus the imposition of a secondary task may have the effect of maintaining (or even improving) performance on the primary

task. It must, however, be recognized that when the operator narrows his limits, his workload has effectively increased, and measures such as RMS error and time-on-target may not indicate directly the extent of operator workload.

DISCUSSION

As the studies above indicate, behavior of the simulated operator is quite consistent with that of a human operator under similar task conditions. The studies were chosen primarily to demonstrate the flexibility of the model, and do not represent an exhaustive exploration of all possible condition combinations. They do, however, show the ease with which the properties of task demands can be modified, since each study required only minor changes in input parameters once the basic simulation was completed. Extended to more complex systems, such rapid and inexpensive task modifications would allow a system to be evaluated for a variety of task and operator characteristics. Sensitivity of such an evaluation to potential performance decrements would be greatly enhanced.

Output from the HOS model and its analysis program (HODAC) can be used to examine operator workload from several viewpoints. Measures of workload commonly obtained in multi-task situations include percentage of available time occupied by task activity (percent workload) and "reserve capacity" measures determined from time available for performing a secondary task. These measures and others can be obtained from the HODAC timeline of operator activity and its "snapshots" of system and operator status at specified time intervals. As the studies above suggest, however, these traditional measures have potentially fatal deficiencies as indicators of task demand, and are difficult to interpret correctly without additional knowledge about task difficulty and about the task conditions under which they were obtained. The strength of the HOS approach for evaluation of multi-task systems lies in its control over operator characteristics and task demand parameters. By adding artificial tasks to an existing task structure, and by systematically varying frequency and duration of artificial task interrupts, potential overload problems can be more clearly identified and understood. When alternative crewstations or competing function allocations are compared, control over task parameters allows precise duplication of operator response and of the pattern of task interrupts within both alternatives. Comparison of simulation outcomes across alternatives under such control provides statements about relative operator loading that are more meaningful than those obtained when task demands within alternatives are not so clearly specified.

Such measures of relative workload are of great value for systems design and evaluation in distinguishing among design alternatives. They are of limited utility, however, as indicators of the absolute workload imposed by system tasks. To assign a unique figure to a set of tasks that reflects the workload requires a comprehension of human information processing and psychomotor task performance far beyond that currently attained. Such a single absolute measure must be completely independent of the content and structure of the tasks on which it is computed, and must therefore be linked to some theoretical unidimensional mechanism underlying all human task performance. It is questionable that such a concept of absolute workload can exist; even if such a metric were available, it is likely to be of no more value to the system designer than coarser measures of workload that point out the location and probable cause of overload problems within the mission timeline.

Measures currently obtainable from HOS, like those from other approaches, are primarily relative loading indicators. The HOS model, with its underlying micromodels, may have greater potential as a source of generalizable workload measures than approaches using a more macroscopic structure. For example, sets of tasks may be identifiable in some unique fashion in terms of the pattern of frequencies of use of each micromodel and by the matrix of transitions between micromodels. Clusters of patterns could define families of task sets, by which apparently dissimilar task sequences could be shown to impose similar loading demands on an operator. The development of additional workload measures from HOS is both possible and desirable, but will require extensive exploration.

The HOS model, in its present form, does not provide for decrements in performance as a result of sustained operator effort over time. Significant fatigue effects are seldom encountered in most short-term system evaluations, and the quality of simulation for most applications of HOS is generally unaffected by the lack of a fatigue micromodel. For long-term performance and for conditions of very high sustained workload, it would be desirable for the cost in operator effort to be reflected in simulated performance. A conceptual approach to the modeling of fatigue for HOS is presented elsewhere (24).

CONCLUSIONS

The HOS system offers considerable potential as a workload-assessment technique for application during the midrange of system development and as an adjunct to analytic prediction methods during early system development. Studies support the applicability of the model to multiple-task situations, with control of task parameters allowing more precise estimates of operator loading. Outcomes of the HOS simulation studies point to critical deficiencies for many workload measures currently in use. Further development is required to refine and expand HOS workload indices and to define procedures for empirical applications of the model. The development of a fatigue micromodel is recommended to support the simulation of extended high-workload task conditions.

REFERENCES

1. Lovesey, E. J. The development of aircraft instruments. In: Visual Presentation of Cockpit Information Including Special Devices Used for Particular Conditions of Flying. NATO/AGARD Conference Proceedings CP-201, November 1976.

2. Schiflett, S. Operator workload: An annotated bibliography. Report No. SY-257R-76. Naval Air Test Center, Patuxent River, Maryland, February 1977. (Limited Distribution)

3. Jahns, D. W. A concept of operator workload on manual vehicle operations. Report No. 14. Forschung-institut fur Anthropotechnick, Meckenheim, FRG, December 1973.

4. Jahns, D. W. Operator workload: What is it and how should it be measured. In: Cross, K.D. & McGrath, J. (eds.). Crew Systems Design. Anacapa Sciences, Inc., Santa Barbara, California, July 1973.

5. Rolfe, J. M. The measurement of human response in man-vehicle control situations. In: Sheridan, T. & Johannsen, B. (eds.). Monitoring Behavior and Supervisory Control. Plenum Press (New York), 1976, pp. 87-93.

6. Howitt, J. S. Flight deck workload studies in civil transport aircraft. In: Measurement of Aircrew Performance - The Flight Deck Workload and Its Relation to Pilot Performance. NATO/AGARD Conference Proceedings CP-56, 1969.

7. Siegel, A. & Wolf, J. A technique for evaluating man-machine systems design. Human Factors, 1961-3, pp. 18-27.

8. Siegel, A. & Wolf, J. Man-Machine Simulation Models - Psychosocial and Performance Interaction. Wiley (New York), 1969.

9. Dickey, L. R. Flight deck certification computer programs: Cockpit crew workload. Report No. D6-29906-3. Boeing Aerospace Company, Seattle, Washington, 1969.

10. Whitmore, D. C. & Parks, D. L. Computer-Aided Function Allocation-Evaluation System (CAFES). Phase IV, Final Report, Volume I (No. D180-18433-1). Boeing Aerospace Company, Seattle, Washington, December 1974.

11. Linton, P., Jahns, D. W. & Chatelier, P. Operator workload assessment model: An evaluation of a VF/VA-V/STOL system. Presented at NATO/AGARD Aerospace Medical Panel Specialists Meeting, Koln, FRG, 18-22 April 1977.

12. Pritsker, A. A. B., Wortman, D. B., Seum, C. S., Chubb, G. P. & Seifert, D. J. SAINT: Volume I. Systems Analysis of Integrated Network Tasks. AMRL-TR-73-126. Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio. April 1974.

13. Murphy, J. V. & Gurman, B.S. The integrated cockpit procedure for identifying control and display requirements of aircraft in advanced time periods. In: Guidance and Control Displays. NATO/AGARD Conference Proceedings CP-96, 1972.

14. Belcher, J. J. A technique for assessing operability/effectiveness of control-display systems. In: Cross, K. D. & McGrath, J. (eds.). Crew System Design. Anacapa Sciences, Inc., Santa Barbara, California, July 1973.

15. Wingert, J. W. Function-interlace modifications to analytic workload prediction. In: Cross, K.D. & McGrath, J. (eds.). Crew System Design. Anacapa Sciences, Inc., Santa Barbara, California, July 1973.

16. Stackhouse, S. P. The measurement of pilot workload in manual control systems. Report No. F0398- FR1 Honeywell, Inc., Minneapolis, Minnesota, January 1976.

17. Spyker, D., Stackhouse, S. P., Khalafalla, A. & McClane, R. Development of techniques for measuring pilot workload. NASA CR-1888. National Aeronautics & Space Administration, Washington, D.C., 1971.

18. Johannsen, G., Pfendler, C. & Stein. Human performance and workload in simulated landing approaches with autopilot failures. In: Sheridan, T. & Johannsen, G. (eds.). Monitoring Behavior and Supervisory Control. Plenum Press (New York), 1976, pp. 87-93.

19. Smit, J. Pilot workload analysys based upon in-flight physiological measurements and task analysis methods. In: Sheridan, T. & Johannsen, G. (eds.). Monitoring Behavior and Supervisory Control. Plenum Press (New York), 1976, pp. 119-124.

20. Wherry, R. J., Jr. The Human Operator Simulator: HOS. In: Sheridan, T. & Johannsen, G. (eds). Monitoring Behavior and Supervisory Control. Plenum Press (New York), 1976, pp. 283-293.

21. Strieb, M. I. The Human Operator Simulator. Volume VI: Simulation description. TR 1181-B. Analytics, Inc., Willow Grove, Pennsylvania, December 1975.

22. Strieb, M. I., Glenn, F. A., Fisher, C. & Fitts, L. The Human Operator Simulator. Volume VII: LAMPS Air Tactical Officer Simulation, TR 1200. Analytics, Inc., Willow Grove, Pennsylvania, November 1976.

23. Kahnemann, D. Attention and Effort. Prentice Hall (Englewood Cliffs, N. J.), 1963.

24. Strieb, M. I., Wherry, R. J., Jr. & Glenn, F. A. The Human Operator Simulator: Applications to assessment of operator loading. TR 1233-A. Analytics, Inc., Willow Grove, Pennsylvania, 1977. (In prep.)

DISCUSSION

DOERING: (Germany)
You mentioned that it is possible with HOS to simulate the human operator and the machine in a man-machine system. Are the interactions with the mission environment or with other systems also simulated?

LANE: (United States)
Interactions with other systems are not simulated, but the influence of the physical environment on the operator is taken into account by means of different operator status.

DOETSCH: (Germany)
How can one place confidence in HOS, considering previous failures with the "Paper Pilot," etc., when applied to such ambitious, complex tasks? I am very pessimistic!

LANE: (United States)
HOS is a different kind of model than "Paper Pilot" and other control theory models which deal with the operator only as a rather inflexible servomechanism. The weakness of control theory models is that only a fraction of an operator's job is psychomotor control of a system; the majority is cognitive, including the processing of feedback from the system and the generation of outputs to the system. Servo models which treat the operator as a transfer function calculator, even though they may be very good models, simply are not flexible enough or complete enough to model the total operator job. They don't account for cognitive functioning or compensate for the quite different strategies used by operators in achieving essentially identical control outputs. Considering the differences between HOS and previous approaches and the capability thus far demonstrated by HOS, we are still optimistic.

BRICTSON: (United States)
What strategies did you use to determine operator internal performance criteria for both simple and complex tasks as well as overall mission completion?

LANE: (United States)
HOS allows the user to set the values for what the "operator" will consider to be satisfactory performance. For example, the operator can be instructed to respond to altitude variations only when they are more than 50 feet from a specified altitude value, and he will not make control corrections so long as readings are in that range. This "limits" value could be set at any other figure desired, and it can be changed dynamically as the mission progresses. For some types of missions, there are relatively clear criteria of overall mission success, such as the touchdown point or arresting wire engaged on a carrier landing. It would be necessary to trace the carrier approach back, as you have done, to determine the allowable envelope or limits of altitude at the quarter-mile, half-mile, etc. points during the approach. One can then instruct the HOS operator to use these limits as internal criteria to judge the goodness of his approach, and if he cannot hold these limits, to initiate a waveoff. In other words, you must determine the intermediate checkpoints during a mission to provide to the HOS operator. These can only come from a reasonably good understanding of the process you are attempting to simulate.

NICHOLSON: (United Kingdom)
Another question on strategy. How do you establish steady performance and constant strategy in carrying out the task? Strategy changes during learning. How do you manage this? How do you determine when the operator has learned the task?

LANE: (United States)
We assume the operator is fully trained.

HOLLOWAY: (United States)
Papers given yesterday (Simmons; Gregoire) indicated that pilots flying standard missions were not following standard procedures for systems operations, although they reported that they were using such procedures. Are you disturbed by the nature of the standard task for operating a system?

LANE: (United States)
The HOS system is for use during mid-range systems design when performance on simulators is not available. HOS is not intended to deal with the problem of individual differences in individual operators. Furthermore, I'm not sure that pilots know exactly how they perform a task.

SANDERS: (Netherlands)
I have two questions. First, do you endorse the statement that HOS cannot be better than current human performance models? Second, how did you validate this?

LANE: (United States)
With regard to models, I suppose I accept your statement in part. I think the interactive nature of this model may compensate for the lack of sophistication of some of the micro-models. Also it may be that highly accurate and sophisticated modelling of each function may be necessary to achieve satisfactory overall accuracy. Validation of HOS was achieved by comparing performance on HOS with operators' analyses of time and accuracy dimensions in a range of simple tasks.

MALECKI: (United States)
In your presentation, you made reference to measures of absolute workload. Would you please elaborate by defining absolute workload and giving illustrative measures?

LANE:
(United States)

As the paper suggests, most measures of workload must be considered relative measures. While they might enable comparison of two ways of accomplishing the same set of tasks, they are not useful on other than that comparative basis, that is, they do not address the workload of a task set. An absolute workload measure, if such a thing exists, would allow a statement about the operator loading associated with accomplishing a fixed task assignment in a given time period which could be compared on the basis of a common metric with the workload imposed by a different set of tasks. I'm personally slightly pessimistic about ever acquiring measures which allow such statements without qualifications, but it should be possible to come closer than we are now. It is unclear exactly what the infrastructure of such a measure would look like, but I'm sure that it would require much better understanding of human information processing capabilities than currently exist, and would likely involve some way of grouping task activities according to the cognitive functions required for their execution. For example, in HOS, it is possible to keep track of how many times each micromodel is assessed in the performance of a task, and thus, for a "standard" task, to compute a profile of the types of cognitive and psychomotor activities involved. It may be possible to group tasks with similar activity profiles into task families which could be used as higher level building blocks for quantifying "absolute" workload. I'm not at all sure how one would go about doing this.

OPERATOR WORKLOAD ASSESSMENT MODEL: AN EVALUATION OF A VF/VA-V/STOL SYSTEM

by
Mr. Paul M. Linton
Crew Systems Department
NAVAL AIR DEVELOPMENT CENTER
Warminster, Pa. 18974

Mr. Dieter W. Jahns
BMAD Avionics-Crew Systems Tech. Staff
THE BOEING COMPANY
Seattle, Wash. 98124

and

CDR. Paul R. Chatelier, MSC, USN
Research and Technology Group (Code AIR 340F)
NAVAL AIR SYSTEMS COMMAND
Washington, D.C. 20060

SUMMARY

The term "operator workload" generally refers to a concept used in evaluating the extent to which a human operator is occupied with meeting system demands. The evaluation and estimation of workload has become increasingly important in the development and operation of advanced aircraft systems as traditional human performance measurement techniques have had to be augmented with other measures to reflect the changing role of the human operator in complex system operations.

By systematically describing the steps undertaken to estimate the workload in a conceptual fighter/attack V/STOL aircraft, meeting Navy mission requirements, this paper shows that while a single crewmember can probably manage the V/STOL in its primary mission phases, additional and refined workload assessment evaluations will be required to conclusively settle the issue for all aspects of deployment. In addition, it is suggested that an integrative concept of operator workload be developed to more accurately reflect the psychological/physiological demands made of the operator in the course of system operations.

INTRODUCTION

The push-and-pull between advanced technological capability and advanced aerospace system performance requirements over the past several decades has vastly changed the role and tasks of the crew operating and controlling these systems. In addition, increasing awareness of our limited resources (primarily in terms of money and energy) and environmental considerations have broadened the factors which must be considered in the development and use of aircraft systems. In the commercial aviation sector, for example, noise abatement requirements have changed take-off and landing procedures in ways which were not fully considered when a large proportion of the aircraft now in service were originally developed. Both equipment and flight crews are being asked to adapt to these new requirements. Similarly, potential gains in operational flexibility and system effectiveness realizable from V/STOL technology applications to fighter/attack (VF/VA) aircraft are continually being studied and incorporated by the U.S. Navy in the development of advanced, manned, airborne weapon systems.

By utilizing new technology in automatic control and avionics hardware and software, among others, increased automation has changed the aircrew's role in system operations from that of a generator and transmitter of physical energy to that of a decision-making monitor with supervisory control. Concomitantly the measurement of the human performance contribution to system operations is also changing. While traditionally it was believed sufficient to quantify the relationship between inputs to the human operator and his overt response to them in affecting a required change in the system, it is now realized that overt responding alone is not sufficient to describe the complex relationships between man and machine. Increased automation may have decreased the number of overt responses the aircraft crew may be required to make, but increased system capabilities may have disproportionately reduced the time available to make those remaining responses and/or added new monitoring tasks the effects of which require different measurement techniques.

This realization that the introduction of automation does not necessarily reduce the involvement of the crew in aircraft operations, but only changes it (Ref. 1), has recently led to renewed interest in and diversification of concepts, models, and assessment techniques for what is commonly known as "operator workload" (Ref. 2). As Rolfe (Ref. 3) puts it: "The definition of workload is more often than not an operational one; it is related to what the systems engineer, physiologist, ergonomist [human factors engineer], or flight medical officer requires it to be." It is the intent of this paper to demonstrate how the systems engineer currently can address the issue of operator workload in the course of developing a complex, airborne system and what research and development efforts are now ongoing and needed in the future, to enhance the state-of-the-art in comprehensive operator workload assessment. The conceptual VF/VA-V/STOL aircraft being studied by the U.S. Navy (both in-house and under contracts to industry) is a "representative solution" to current operational requirements and will serve as a vehicle for illustrating the problem of workload assessment very early in the development cycle of a new system. The workload assessment model used is an integral part of a comprehensive human engineering tool (Computer Aided Function-Allocation and Evaluation System, CAFES) being developed under Naval Air Development Center contract by Boeing (Ref. 4).

SYSTEMS ENGINEERING CONCEPT OF WORKLOAD

One definition of operator workload which has found widespread use in certification programs for the operability of both military and civilian flight decks in the USA has its foundation in motion and time study. As used by many systems engineers, workload is the extent to which an operator is occupied by a task relative to the time that is available for accomplishing the task. This definition is illustrated in Figure 1 which shows the following case: An operator has to complete a particular switching task within a specific time period T_A during which he has to obtain decision information, determine his action, and switch some equipment. Information is gathered by finding and reading a display which takes T_1 amount of time. The switching necessitates finding the (hopefully) correct switch and activating it, requiring an amount of time T_2. Now, workload becomes the ratio of required time ($T_1 + T_2$) to that available T_A, i.e., expressed as percentage figure:

$$\text{percent workload for task} = \frac{T_1 + T_2}{T_A} \times 100 \qquad (1)$$

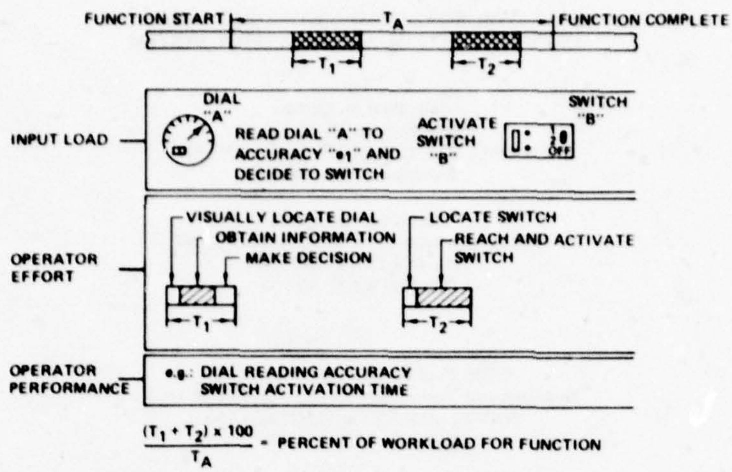


Figure 1. Systems Engineering Concept of Workload

Note that performance accuracy is not directly considered in this definition, but it is generally assumed that the performance being timed meets preestablished criteria. The time available for task execution by the operator is usually determined by sequential system functions over which the operator has little, or no, control.

As will be amplified later, procedures involved in the systems engineering approach to workload evaluation rely heavily on detailed task analyses augmented by techniques derived from motion-and-time engineering. In addition, since this approach is primarily used during system development, it involves a broad concept of workload assessment which nevertheless has a very specific mission/equipment/operator relationship as its foundation. Any partitioning of this triad may serve the interest of the research community, but compromises the data needed by the systems engineer in configuring a system with an acceptable mix of automated and manual functions within available resources.

The following section will describe an application and the techniques used in the systems engineering approach to workload assessment.

WORKLOAD ASSESSMENT IN A CONCEPTUAL VF/VA-V/STOL AIRCRAFT

At the request of the Naval Weapons Center (NWC), China Lake, California, the Human Factors Engineering Division, Naval Air Development Center conducted a crew loading analysis of the NWC conceptual VF/VA-V/STOL aircraft. The purpose of the study was to make a preliminary assessment on the ability of a single crewmember to effectively execute selected VF/VA-V/STOL missions. Since the aircraft in question was only a "representative solution", i.e., an aircraft concept existing only at a gross level of detail, it was recognized that the results of the loading analysis would only be valid within the context of assumptions made relative to the mission scenarios, cockpit configuration, and task definition and sequencing defined for the study.

A step-wise approach to the crew loading evaluation was adopted. The eight analytical tasks and interim documentation comprising the total effort are depicted in Figure 2. The following brief discussion of this application of the CAFES Workload Assessment Model is structured to correspond directly with the eight tasks defined in the figure.

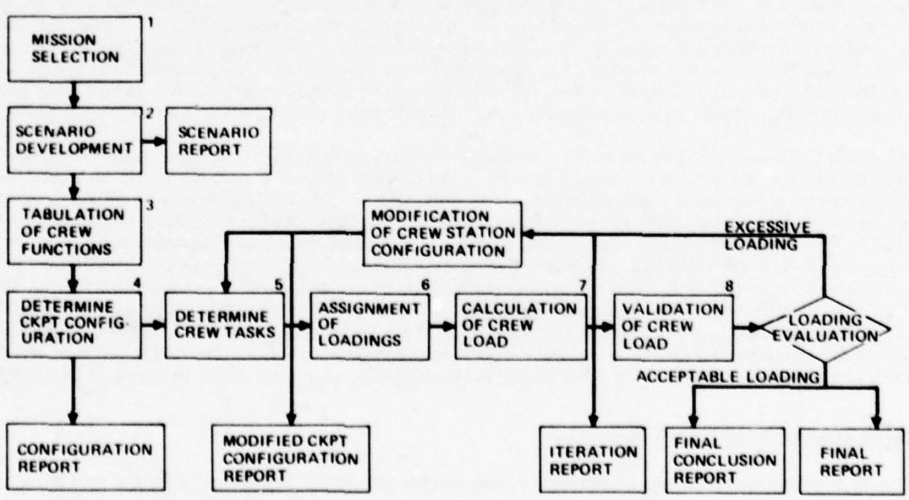


Figure 2. VF/VA-V/STOL Crew Loading Analysis and Documentation

1. Mission Selection

The initial step in the crew loading analysis was the selection of representative fighter/attack (VF/VA)-V/STOL missions. Numerous VF and VA reports were examined during a concentrated familiarization and literature search period; particular attention was paid to anticipated missions and assumed avionics suites. Naval Weapons Center Memorandum 4008-128, "VFA-V/STOL Preliminary Mission Identification and Functional Analysis" concluded that only four missions -- two fighter and two attack -- were essential to the VF/VA-V/STOL concept. The fighter missions identified were deck launched intercept and air superiority, and the two attack missions were

anti-shipping strike and close air support. It was decided to select one mission from each area in order to examine the full range of VF/VA-V/STOL responsibility. From the fighter series, deck launched intercept was selected over air superiority when a closer examination of the two missions indicated that deck launched intercept required more complete pilot-aircraft interaction than did offensive air superiority. Most air superiority missions reviewed were resolved through air-to-air engagements, which are probably more stressful than a typical deck-launched intercept engagement, but require less of the pilot-avionic interaction. Once the air-to-air battle is engaged, the pilot must keep his adversary in sight and maneuver for the tactically advantageous position. His head is out of the cockpit, and his primary concern is flight control.

In the attack series, close air support was selected over anti-shipping strike since the former presented the more demanding tasks for the operator. Target acquisition and identification, for example, were judged to be more difficult in close air support. In addition, enemy defenses are usually unpredictable during these missions, and communications between the pilot and his forward air controller demand considerable attention.

2. Scenario Development

The Close Air Support (CAS) and the Deck Launched Intercept (DLI) scenarios are presented in narrative form in Appendix A. They were developed specifically for this crew loading analysis after a number of similar missions and VF/VA-V/STOL requirements were researched, and serve two basic purposes. First, they describe in some detail realizable missions which are sufficiently comprehensive to allow further human factors and systems analysis, and second, they provide a general appreciation of the avionics and system concepts attributable to, and expected of, the VF/VA-V/STOL aircraft.

3. Tabulation of Crew Functions

Functional Flow Block Diagrams (FFBD) had to be generated prior to defining the crew functions. These flow diagrams present in rough form all those specific events which must be satisfactorily completed to accomplish the mission.

FFBD's were only a tool to assist in the task analysis. They were not intended to be a final product and did not contain every function that may have to be executed during the mission. Some functions did not appear in the flow because they were continuous in nature and obviously machine tasks. The MONITOR ELECTRONIC ENVIRONMENT function, for example, must be allocated to an electromagnetic sensor and need not appear discretely in the function flow. The human operator has a complementary function (SCAN THREAT DETECTION AND WARNING DISPLAY), which he executes on an as-required basis. This operator function appears in the flow periodically either when the pilot is making a routine instrument check or when the function is an integral part of the mission situation. Some other continuous functions that are not specifically identified in the flow diagrams are:

MAINTAIN FLIGHT CONTROL
COMPUTE GEOGRAPHIC POSITION
DISPLAY GEOGRAPHIC POSITION
PROVIDE SELF IDENTIFICATION
MONITOR FLIGHT INSTRUMENTATION

Conducting a crew loading analysis on the complete mission was obviously very inefficient. Mission phases such as pre-flight, climb to altitude, etc., presented no workload problem and were not examined in any detail beyond the scenarios. The only mission phases carried to the FFBD level were those where the peak loading was expected, i.e., pre-attack, attack, and escape. FFBD's for the selected phases of the CAS and the DLI missions are presented in Ref. 5.

4. Cockpit Configuration

The arrangement and selection of the controls and displays assumed for the VF/VA-V/STOL cockpit provide the basis for translating the functions identified for mission accomplishment into the specific tasks and task times required for the loading analysis. A crew station combining the features of standard control/display technology, advanced integrated instrumentation concepts, and some original design was developed and fabricated for this purpose. Figure 3 presents the overall appearance of this cockpit. Figure 4 shows programmable displays, implemented through slides and video tape, on the four panel display units. The Head Up Display (HUD) format is superimposed upon an external screen in front of the cockpit. This projection is positioned so it falls in the same place in the pilot's field-of-view that would be occupied by an actual head-up display.

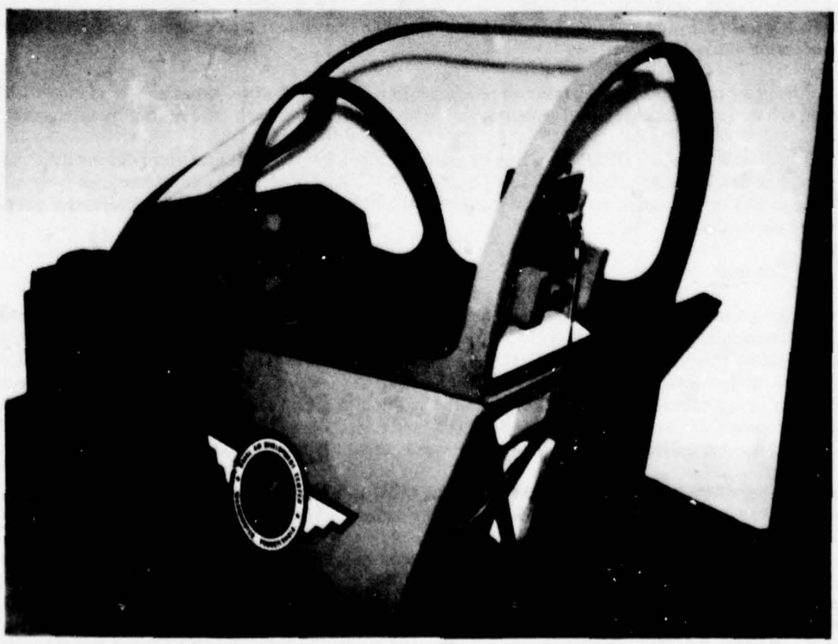


Figure 3. Overall Cockpit Configuration

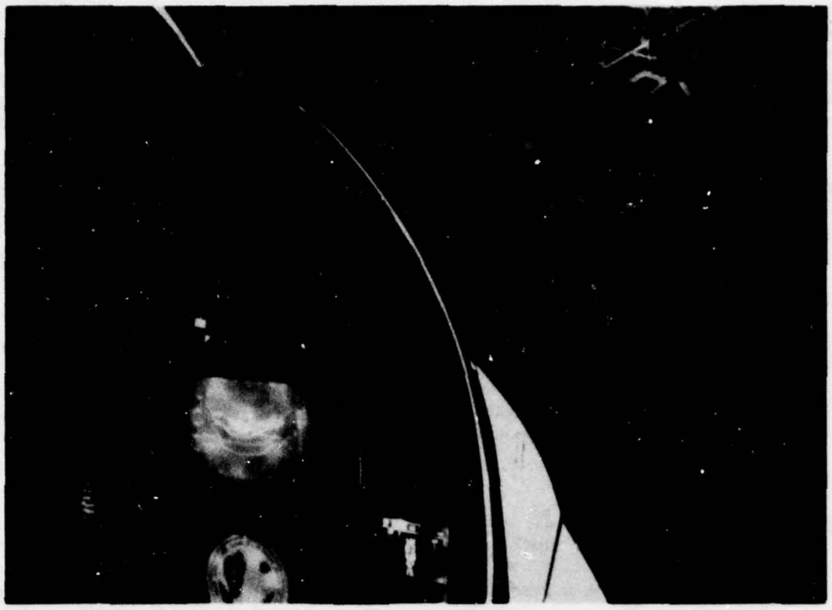


Figure 4. Representative Programmable Display Formats

The detailed control and display configurations are not contained in this paper due to space considerations. A thorough discussion of each display and panel within the cockpit can be found in Ref. 5. Collectively, the controls and displays selected for the cockpit encompass everything necessary for the information exchange required to accomplish the mission scenarios. Though reference was made on a continuing basis to current practice and proposed developments for subsystems (e.g., TV, LASER, RADAR, etc.), no attempt was made to specify exact electronic interconnections and/or interfaces. Therefore, some switching functions may eventually be consolidated and others revised and/or separated. However, here again, the designs presented are felt to be sufficiently representative to permit the completion of valid task definitions.

5. Tabulation of Crew Tasks

This part of the analysis was relatively straightforward for the analyst familiar with the functional flow block diagrams and the VF/VA-V/STOL cockpit arrangement. The analyst simply identified what specific display surfaces and controls were utilized to physically carry out the function in a particular block. He also defined how the control was to be operated (hands, voice activated, etc.) and whether operator cognitive processes were required. This determination had to be made for each function to be performed during the conduct of the mission. Complete task analyses for the CAS and DLI missions are contained in Ref. 5.

6. Assignment of Task Loadings

Although called out as a separate step, this process was really done at the same time the crew tasks were being tabulated. The point of interest here is how the task times and operator channel loadings were assigned. Obviously, the loading computations and hence the conclusions of the study are dependent upon the times and operator channel loads assigned. In this study, task times were assigned either analytically, experimentally, or by estimation. In some cases, combinations of these methods were required. In addition, every event was considered in context before a time was assigned. In many cases two events, each containing the same task, were assigned different durations depending upon their position in the mission. For example, an engine instrument check can be performed more leisurely during the cruise phase than it can during a dive to target.

Discrete switch depressions and knob turning tasks were assigned reach and actuation times in accordance with human factors handbook data. Extensive experimentation has been performed in this area, and the handbook data is considered accurate and valid.

Less standard tasks such as RELEASE ORDNANCE were assigned times by experimental observations utilizing the VF/VA-V/STOL cockpit and stopwatches. Extremely variable tasks such as CONDUCT VISUAL SEARCH could be assigned only estimated times on the basis of previous experience and prior studies. In all cases, estimated times were assigned from the conservative portion of the range considered reasonable for the particular task.

7. Calculation of Crew Loading

A very significant question in any workload analysis is how a quantitative measure of crew loading is to be determined. In other words, what data analysis techniques should be selected? As long as the scenario is reasonable and the FFBD is complete, the identification of operator tasks should be reasonably "objective". However, the manner in which the data is to be examined, rather than the composition of the data, will have the greater influence on results. Workload can be assessed through strictly analytical techniques, strictly experimental techniques, or any combination of the two. In addition, sophistication may vary widely within each technique. Workload evaluation by experimental means, for example, can vary from the "untrained operator cardboard mockup stop watch technique" to the opposite extreme of "closed loop simulation operational cockpits and actual flight dynamics".

After evaluating a number of factors, the technique considered most appropriate for this VF/VA-V/STOL analysis was the Statistical Workload Assessment Model (SWAM) of the CAFES (Computer Aided Function Allocation and Evaluation System) (Ref. 4).

SWAM is designed to compute workload percentages relative to combinations of procedures, events, and operator tasks associated with a specified mission being flown in a specified aircraft. The underlying structure of the computer program is a deterministic time-line analysis of the extent to which the operator is occupied with meeting system demands. To facilitate this analysis, nine anatomical/physiological channels have been defined which the operator uses in performing tasks assigned to him. These channels are: (1) internal vision, i.e., internal to the crew station, (2) external vision, i.e., external to the aircraft, (3) right hand, (4) left hand, (5) right foot, (6) left foot, (7) verbal, i.e., speaking, (8) auditory, and (9) cognition.

As was discussed above, a significant effort in preparing inputs to SWAM consists of determining for every task to be performed (a) which channel(s) will be used in performing the task and (b) for what percentage of task duration time each channel will be used. For an example, refer back to the switching-task shown in Figure 1. The operator will use his eyes (internal vision) to locate the dial, obtain information, and locate the switch; he will use his right hand to reach and activate the switch; and he will use his cognition to process the information received and make a decision. Although the relative time allocation may be as shown in Figure 1, it is assumed that the internal vision channel will be used for 100% of the total task time ($T_1 + T_2$ in Figure 1). The right hand channel is occupied 45% of the total task time and the cognitive channel is occupied 40% of the time required to perform the task.

As another input to SWAM, the mission or mission phase, is structured into a series of events (in this case using the FFBD's discussed above). An event is a single, discrete operator activity, and each event consists of a task label or code, a start time relative to mission start, and the event duration (i.e., $T_1 + T_2$ in the example). This list of sequential events, including the channel allocation data, forms the base upon which SWAM performs statistical analysis. SWAM examines the mission as a series of arbitrary, preselected time segments. For the VF/VA-V/STOL case, a six-second segment was chosen, and SWAM computed the percent workload imposed upon each anatomical/physiological channel during each sequential segment throughout the mission phase being considered. Returning to the example, the selection of the time segment is equivalent to setting T_A equal to six seconds for all tasks.

Numerous output options are available for SWAM in digital printer and graphical plotter formats including mission scenario, crew workload profile, workload summary statistics, task channel activity, and task lists, among others. Some of these are discussed with the results of the VF/VA-V/STOL analysis below.

8. Validation

The purpose of this step was to assure that assignments of task times made in step 6 were reasonable and proper. As previously mentioned, three methods – analytical, experimental and estimative – were utilized to obtain task times.

The analytical task times used for discrete actions such as switch depressions had already been authoritatively validated, and little question existed regarding their accuracy or applicability. The experimentally assigned task times could not be validated other than by observing and recording inflight performance or by performing more sophisticated simulations. These experimentally derived task times were themselves the averages of a number of observations recorded with varying numbers of subjects. The times were also reviewed by the Crew Systems Technology Department of the Boeing Aerospace Company and found to be reasonable. SWAM has previously been validated by Boeing for both military and civil applications.

The time estimates assigned to loosely defined tasks such as conducting a visual search are just that – estimates. Obviously, no "correct" time can be assigned to these tasks. Often a pilot will visually detect his target within a few seconds after commencing the search; on other occasions, he may search for 15 or 20 seconds and still fail to make a detection. So many variables influence this type of task that pilot comments, past experience, and the philosophy of keeping the estimates conservative provided the guidelines for time assignments to variable length tasks.

9. Results and Loading Evaluation

There is insufficient space in the pages allocated for this paper to present all the results obtained from the application of SWAM; these are delineated in Ref. 5. The channel activity summary plots for the close-air-support (CAS) and deck-launched-intercept (DLI) mission phases are shown in Figures 5 and 6 respectively. These plots indicate to what extent each of the operator channels was active on the average (plus one standard deviation) throughout each respective mission phase. Summaries are also provided for overall vision, motor (hands and feet) and communication (verbal, auditory) activities, was well as a total average workload across all channels throughout the mission phase. As can be seen by comparing the data of each figure, the missions are practically equal in average workload with the CAS mission slightly higher than the DLI mission in the majority of channel categories.

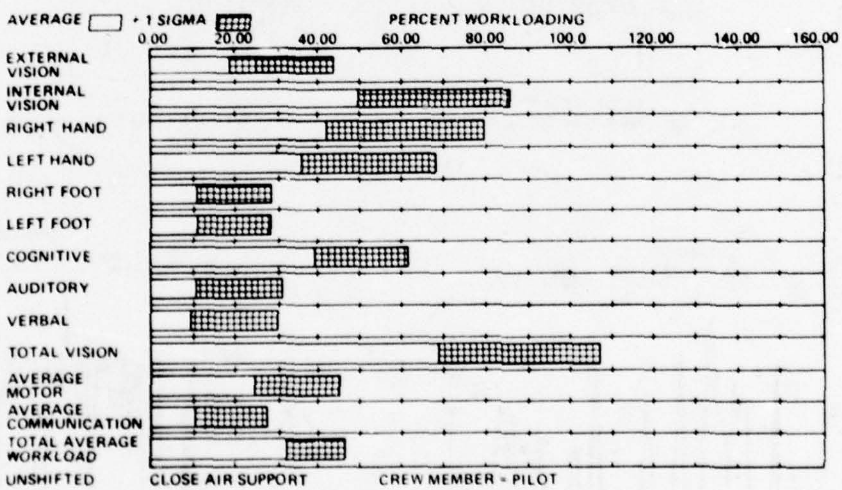


Figure 5. Channel Activity Summary Plot for the Close Air Support Mission Phase

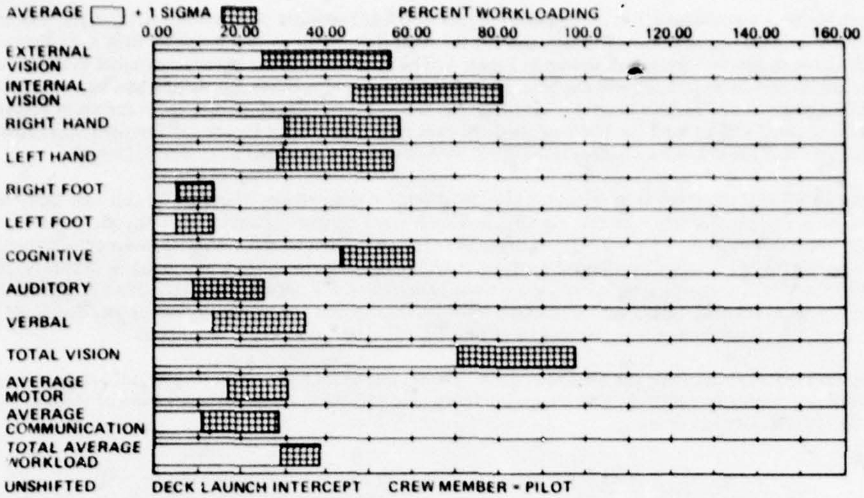


Figure 6. Channel Activity Summary Plot for the Deck Launched Intercept Mission Phase

Plots of percent workload vs. time for selected channel summaries are shown in Figures 7A through 8C. These plots are most useful for quickly locating at what point(s) in the mission and in which operator channel overload conditions occur. Theoretically, an overload occurs on a channel when its utilization exceeds 100% for an extended period of time beyond the "resolvable time" set by segment size (i.e., six seconds in this case). Anytime this situation occurs (see Figure 7C Total Vision Plot), the tasks being executed during the interval (or segment) in question must be reexamined to determine how (or whether) they can be separated time wise. If one or more of these tasks cannot be shifted to elsewhere in the function sequence to relieve the overload, a re-allocation of functions between man and machine may be required. For the spike in Figure 7C, a shifting of the associated task to a point of lower visual load was possible. Periods in which workload exceeds 80% also must be noted. Although no actual task shifting may be required in this case, much of the workload literature indicates that a continuous workload (i.e., extending over several minutes) of 80% or greater is the approximate point at which performance may begin to degrade. This is especially the case when tasks may be influenced by additional factors such as fatigue, stress, or situations requiring unusual or infrequently exercised responses.

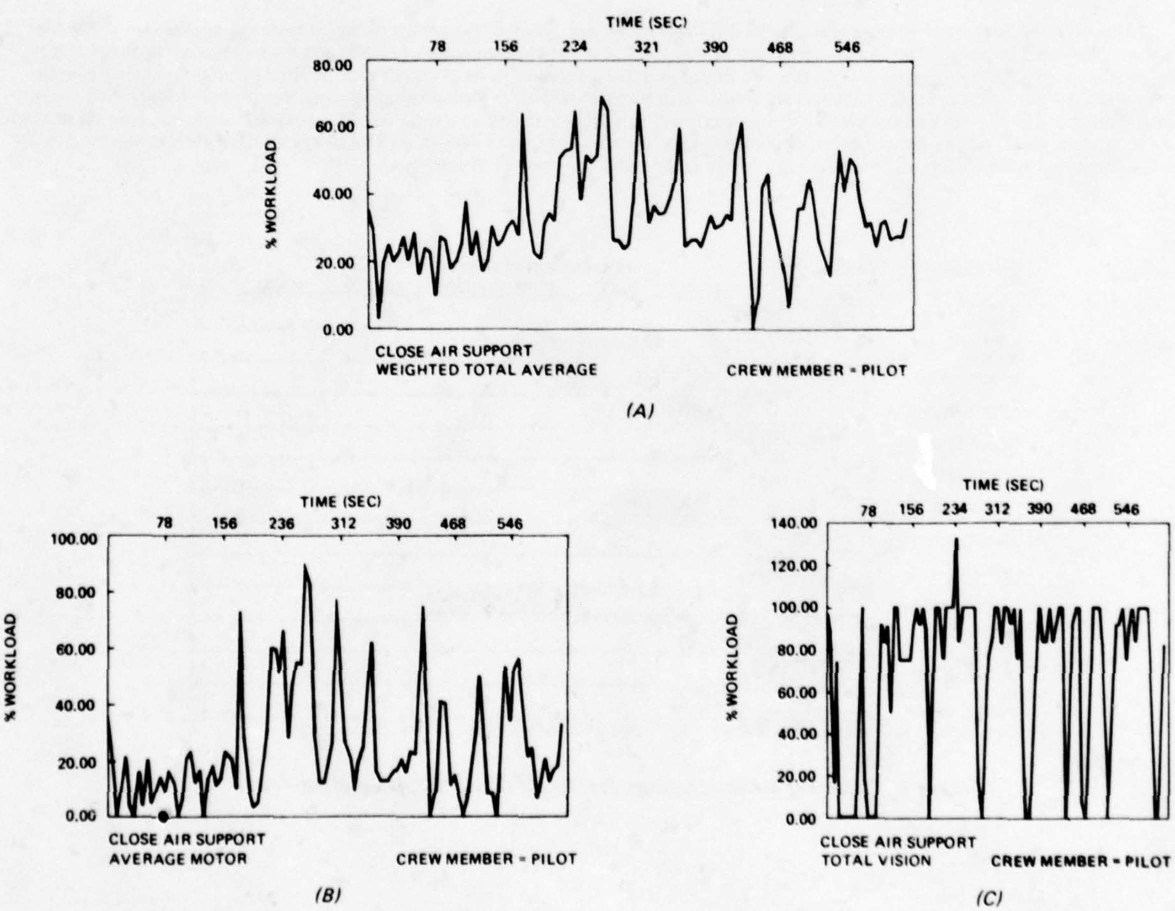


Figure 7. Percent Workload Versus Mission Phase Event Time for CAS

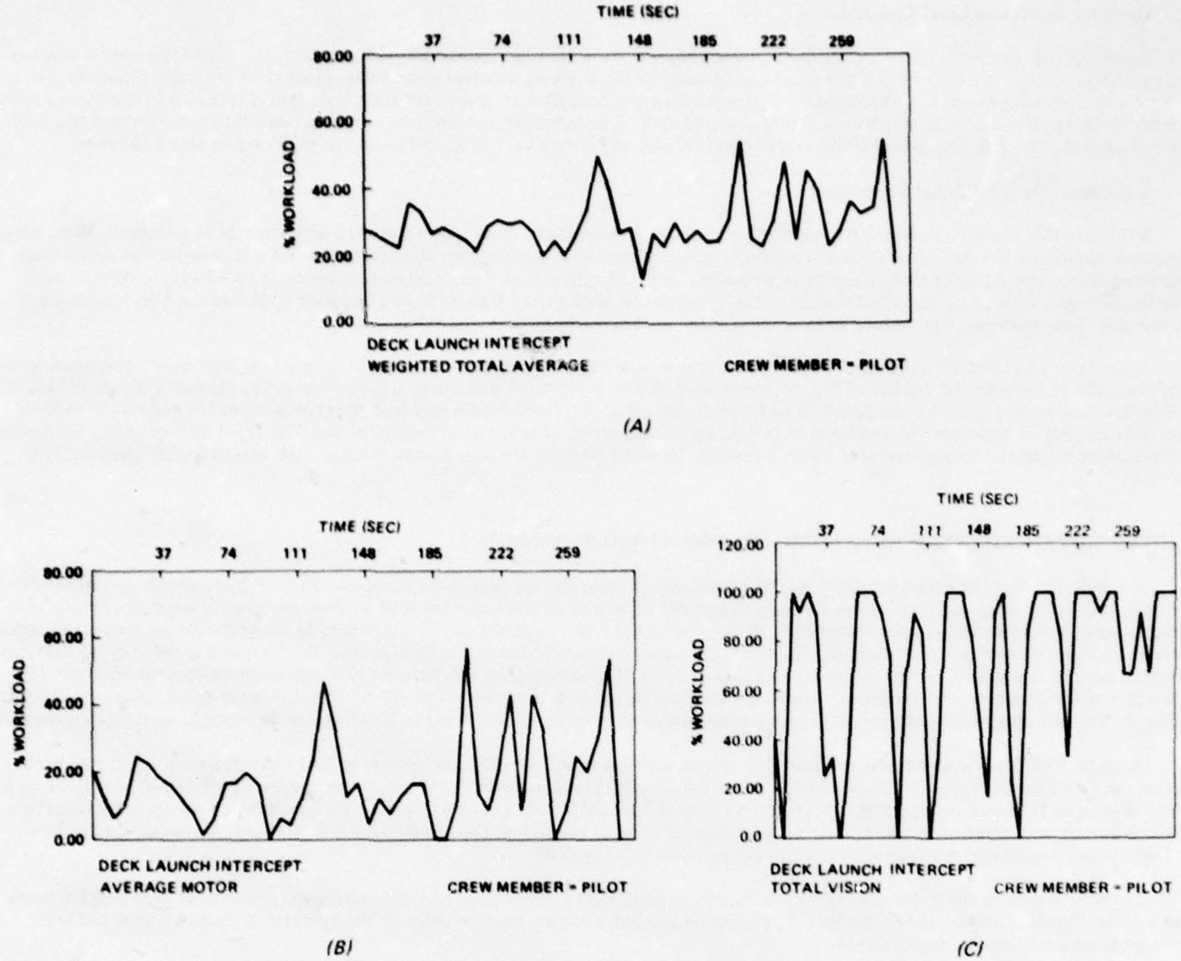


Figure 8. Percent Workload Versus Mission Phase Event Time for DLI

With regard to the present VF/VA-V/STOL analysis, many instances were found to reach 100% workload, but these instances were momentary and thus judged acceptable. No instances were observed during either CAS or DLI which indicated a sustained workload at the 80, 90, or 100 percent levels. Consequently the conclusion was reached that, within the context of the assumptions made, a single operator can manage the VF/VA-V/STOL aircraft during close-air-support and deck-launched-intercept missions.

It should be noted that any workload estimation using SWAM will naturally become more accurate as the particular aircraft system becomes more defined since then more specific data can be submitted for analysis. The present analysis assumed a generalized, futuristic avionics package of the Basic Avionics Subsystem Integration Concept (BASIC) generation. A more precise definition of the controls, displays, and particularly the functional implementation of the VF/VA-V/STOL would allow a more detailed and precise task analysis and subsequent channel activation allocations.

In the same vein, an assessment of the one- or two-man crew question for the VF/VA-V/STOL may be more meaningful if many missions and mission variations were considered. High workload situations not apparent in the CAS and DLI missions should be identified and examined. Peripheral areas such as emergency procedures could also be investigated for high-workload/safety-of-flight implications. For example, two potentially difficult mission phases not examined in this study are vertical take-off and vertical landing. These phases may be very demanding on the pilot in terms of number of required tasks, task precision, and operator stress.

OTHER WORKLOAD ASSESSMENT TECHNIQUES

The validity of any workload technique is dependent upon many factors, most of which fall into two generic categories – validity of inputs and method of analysis. The application of SWAM described above is strictly a statistical analysis of input data obtained from diverse data sources on system and human behavior. Although very useful in bringing all aspects of system operations together into a common time-based framework, SWAM and other motion-and-time techniques for workload assessment cannot by themselves handle the adaptability of the human operator to task situations in real world environments. Thus, while these techniques are system development tools, other techniques are generally used when a more precise estimate of the operator's involvement in task performance is desired. These latter techniques which the systems engineer would likely view as "part-task" techniques to be used in laboratories can be divided into three general classes:

1. Information Processing Techniques

The use of information processing methods to determine operator workload is based on the premise that the human operator possesses a fixed, limited channel capacity and that workload is either the degree to which tasks, additional to a primary one, can be accomplished or the degree to which auxiliary tasks interfere with the accomplishment of the primary task. The underlying concept here is similar to that of motion-and-time techniques, except that the emphasis is shifted toward a quantitative determination of "excess channel capacity" as a measure of operator workload.

2. Operator Activation-Level Techniques

Briefly stated, the "activation" or "arousal" hypothesis maintains that the level of physiological activity within the central nervous system (CNS) is increased both by the mental demands made by the working situation and by the intensity of physical stimuli. In the area of workload assessment, it is hypothesized that input load is a determinant of the activation level that is reflected in the physiological activity of the individual. Thus, physiological measurement techniques are used to correlate changes in heart rate, sinus arrhythmia, EEG, cortical evoked potentials, integrated EMG, respiration rate, and pupillary dilation, as well as others, with various task conditions.

3. Equipment Design-Implicit Techniques

Studies in this category maintain that changes in operator performance are directly associated with changes in workload. Thus, performance improvements that can be related to changes in equipment or procedures are also assumed to indicate a reduction in operator workload. Subjective rating of task difficulty is generally combined with operator performance measures to deduce that operator workload has changed. Unfortunately, most studies of this type include only a brief statement that operator workload has been changed but do not state how workload was defined or how the conclusion was reached.

It has been proposed by Jahns (Ref. 2) that components of all these techniques be integrated into a comprehensive conceptual framework in order to improve the validity of inputs to motion-and-time based workload assessment models and to provide a structure against which data requirements can be compared. In addition, inclusion of these techniques may lead to refinement of the method of analysis currently employed. For example, inclusion of physiological parameters in terms of an "operator state" channel may be useful for dealing with changes within the human operator which influence the effort he must exert in accomplishing a task within specific performance criteria.

FUTURE ADVANCEMENT OF WORKLOAD ASSESSMENT METHODOLOGY

It is believed that the advancement of workload assessment methodology will occur along two lines of development: a) increased use of computer technology for preparation of analysis inputs and workload computation, and b) incorporation of more applicable and descriptive parameters of workload. Under the first line, techniques have been developed which greatly facilitate the preparation of input data by storing lists of tasks, missions, and equipment configurations applicable to a NASA/Boeing 737 for quick recall in event assembly (Ref. 6) prior to analysis. A related effort is being continued by Boeing within an independent research and development program. The objective of this program is to develop a quantitative index of operator performance capacity over mission time during given vehicle operations. An iterative approach of analysis, experimentation, and concept refinement is being used to broaden the scope of workload assessment.

Both the U.S. Navy and industry are actively engaged in research and development efforts to find more descriptive parameters of workload. One example for ways to evaluate the contribution of the human operator to system operations is provided by the paper on the Navy developed Human Operator Simulator (HOS) presented by LCDR N. Lane at this meeting. Other efforts under Navy sponsorship involve the development of test and evaluation techniques needed to assure that the capability and limitations of weapon system crews are adequately considered in the development of advanced evolving systems.

Taken all together, the above efforts are continually heading in the direction of not only asking the question of how well the operator will meet system/mission requirements (i.e., performance), but also how much it will cost the operator in "blood, sweat and tears" *to meet these performance requirements.*

REFERENCES

1. Edwards, E., Some Aspects of Automation in Civil Transport Aircraft, in Preprints to International Symposium on Monitoring Behavior and Supervisory Control, Berchtesgaden, FRG, March 1976, (Dir. G. Johannsen, FAT, 5309 Meckenheim, FRG) p. 2-11.

2. Jahns, D. W., A Concept of Operator Workload in Manual Vehicle Operations, Research report No. 14, Forschungsinstitut fuer Anthropotechnik, 5309 Meckenheim, FRG, December 1973.

3. Rolfe, J. M., The Measurement of Human Response in Man Vehicle Control Situations, Preprints to Monitoring Behavior and Supervisory Control Symposium, Berchtesgaden, FRG, March 1976, p. 97-107.

4. Whitmore, D. C., and Parks, D. L., Computer Aided Function Allocation - Evaluation System (CAFES), Phase IV, Final Report, Vol. 1 & 2, Boeing Aerospace Company Doc. No. D180-18433-1, -2, The Boeing Company, Seattle, Washington, December 1974.

5. Linton, P. M., VFA-V/STOL Crew Loading Analysis, Report No. NADC-75209-40, Naval Air Development Center, Warminster, Pennsylvania, May 1975.

6. Miller, K. H., Timeline Analysis Program (TLA-1), Final Report, Boeing Commercial Airplane Company Doc. No. D6-42377-5, prepared for National Aeronautics and Space Administration, Langley Research Center (NASA-CR-144942), April 1976.

APPENDIX A

DETAILED SCENARIOS

A.1 CLOSE AIR SUPPORT

SCS-81 (Sea Control Ship) is on station 160 miles southwest of the battle area. At a pre-dawn briefing, VFA-V/STOL squadron pilots are informed that major ground offensives are to be initiated at daybreak. It is anticipated that CAS (Close Air Support) will be required, and two of SCS-81's VFA-V/STOL's have been configured for this mission; anti-personnel and armor piercing munitions have been loaded, the modular CAS peculiar avionics have been installed, and all auxiliary CAS software has been loaded. Routine briefing information is exchanged including weather, anticipated defenses, modes and codes of the day, departure and battle area communication frequencies, and launch times. The two-aircraft flight has been designated Yankee Flight. Detailed pre-flight inspections had been completed earlier, and the pilots quickly move to their aircraft and start engines. System checks are run while that equipment requiring initialization is brought on line. The first VFA-V/STOL is spotted, launch clearance is obtained, and an STO (short takeoff) is executed. Table A-1 contains the mission phases.

Yankee Leader retracts gear and flaps and begins transition to completely conventional flight. A quick glance at his emergency/warning multi-purpose display re-assures him that all systems are functioning properly. He begins a climbing turn to the selected inland heading and, visually spotting his wingman, begins accelerating to maximum endurance airspeed and altitude. He notes on his HSD (Horizontal Situation Display) the electronically generated way points and checkpoints superimposed over the projected map presentation. Switching UHF frequencies, Yankee Leader contacts the Direct Air Support Center for his specific assignment and enroute clearance. He and his wingman are assigned to FAC(A)-Alpha (Airborne Forward Air Controller-Alpha) and are given a vector to Alpha from their coast checkpoint. The outbound cruise is routine. The passive threat detection and warning system is very quiet; the radar is placed in STANDBY, and the prominent coastal checkpoint will be acquired either visually or electro-optically.

TABLE A-1. CLOSE AIR SUPPORT - MISSION PHASES

1.0 PRE-FLIGHT
2.0 LAUNCH
3.0 CLIMB TO ALTITUDE
4.0 RENDEZVOUS
5.0 OUTBOUND CRUISE
6.0 DESCEND
7.0 LOITER
8.0 PRE-ATTACK
9.0 ATTACK
10.0 ESCAPE
11.0 CLIMB TO ALTITUDE
12.0 INBOUND CRUISE
13.0 RENDEZVOUS
14.0 RECOVER
15.0 POST FLIGHT

At dawn the ground feature checkpoint is detected and confirmed electro-optically by LLLTV (Low Light Level Television). Yankee Leader raises FAC(A) Alpha on the pre-assigned secure communication channel and receives his specific mission assignment and a more complete briefing. FAC(A) informs Yankee Flight than an advance enemy force has moved up during the night and will soon be within striking distance of a strategic friendly position. The enemy force consists of many troops, light armored vehicles, and rocket launchers. Recent experience has indicated that the enemy will be equipped with portable IR homing missiles and will not utilize any air support. The allied offensive plans have been negated by the enemy move-up, and preparations are underway for defense against the imminent attack. FAC(A) Alpha provides an updated vector from the coastal checkpoint and assigns Yankee Flight a loiter position. Yankee Leader provides an estimated time of arrival, and as he crosses the coastline, he performs a NAV update, reduces power, and drops to a terrain-following altitude. The penetration route has been chosen to minimize detection. No known enemy radar sites are along the route. No rockets are indicated by the threat detection and warning set. Anticipating his upcoming strikes, Yankee Leader reviews his ordnance load via a tabular display on one of his multi-purpose CRT's (Cathode Ray Tube).

Reaching the loiter point, Yankee Flight re-establishes contact with FAC(A) Alpha for update and target information. Alpha reports that friendly troops are presently pinned down under enemy rocket and mortar shelling. Strikes against enemy artillery could force the enemy to seek cover, which would allow friendly ground troops to move up. FAC (A) Alpha and Yankee Flight establish visual contact, and the FAC(A) reports that he will provide data-linked electronic coordinates of the enemy position. Since the weather is excellent, and the attack will be executed visually, Yankee Leader selects a tactical presentation on his HSD (Horizontal Situation Display) and selects the air-to-ground mode on his HUD. Also, since the enemy is still removed from the friendly forces, Yankee Leader decides on a low-level approach, pop-up, and rocket attack. To take maximum advantage of a possible surprise element, Yankee Red will execute the same attack from a different heading. Yankee Leader designates the reference coordinates on his tactical display to receive automatic steering commands; due to the relatively flat terrain he selects a narrow field of view television display on the VSD (Vertical Situation Display). Verifying and arming his ordnance selection, he coordinates with Yankee Red and initiates his low-level, high-speed dash. Maintaining a precise heading, Yankee Leader detects his target about 6 miles out and makes minor steering corrections. He pops-up, enters the dive, tracks his targets, and launches his ordnance. FAC(A) Alpha confirms that his attack has eliminated two of the enemy rocket launchers. Yankee Red has equal success 10 seconds later. Circling high above the battle area, FAC(A) Alpha reports that the remaining rocket launchers are moving to cover while ground troops and many small vehicles are advancing.

To further weaken the enemy force before close engagement, Yankee Leader and FAC(A) Alpha, in conjunction with the FAC(G) (Forward Air Controller-Ground), decide upon low-level strikes utilizing the aircraft's two 30 mm cannons. Yankee Leader's break-off has removed him from the immediate battle area, but the coordinates utilized for the rocket attack were precise and he should be able to quickly return to the battle area. With the element of surprise gone, Yankee Leader is primarily concerned with encountering the IR missiles. After coordinating with FAC(A) Alpha and Yankee Red, Yankee Leader initiates his return to the battle. Commanding the high resolution electro-optic sensor (extremely NFOV) to slew to designated reference points on his tactical display, Yankee Leader can begin a preliminary search of the target area before visually acquiring it. Assuming that any mobile SAMS and IR missiles are confined to the immediate battle area, Yankee Leader climbs to a better search altitude and reduces airspeed somewhat. He monitors his threat detection and warning set carefully. Selecting fine manual slew on the E/O (electro-optic) sensor, the pilot detects some small vehicular traffic. Yankee Leader increases power and heads directly along the bearing to these targets. Still searching with his E/O sensor he notes what

appears to be troop activity and makes the necessary steering adjustments. Not completely sure of the activity he has detected, he requests FAC(G) to mark the forward edge of the friendly position via LTD (laser target designator). Sweeping the LTD sensor in the direction he believes to be toward the friendly forces, he detects the laser returns and confirms his orientation of the battle area. He is now positive that the activity he has detected with his E/O sensor is indeed the enemy advancement. Yankee Leader drops to a lower altitude and attempts to transition from E/O to visual detection. Flying directly toward the enemy line, Yankee Leader visually acquires the intended target area. He verifies that guns are selected and notes the number of rounds still available. He trains his sight on the intended target area and when he receives the in-range indication he begins a series of short bursts along the enemy forward line. Detecting a prominent ground flash he immediately assumes missile launch, breaks off his attack, and takes advantage of his craft's unique maneuvering capability to avoid the missile. Climbing to a safer altitude he sees Yankee Red completing his attack. Both Yankee Leader and Yankee Red turn for one more attack run. Following this strike, the friendly forces take advantage of the temporary confusion and disorganization to launch a counter-offensive and the engagement becomes too confined to allow any further air support. FAC(A) Alpha releases the strike aircraft and informs the Direct Air Support Center, which releases Yankee Flight for a return to ship. Yankee Flight departs on the pre-established heading, airspeed, and altitude. Yankee Leader suddenly receives a voice synthesized threat warning, confirmed by an indication on his threat warning display, that he is being illuminated by a tracking radar. His threat display presents the source of the emitter, SAM-HIGH, and a recommended action, ATTEMPT JAM. He activates the ECM transmitter, releases chaff, and makes a quick visual search for a possible missle launch or even an in-flight missile. He sees nothing. The previous warning is now replaced by a higher priority warning, "EMITTER CLOSING". Since Yankee Leader cannot see the missile, he instinctively rolls into a very sharp breakoff. Pulling out at low altitude, he finds that he has broken track and elects to dash to the coast at terrain following altitude. Feet wet, Yankee Flight climbs to a greater altitude, selects local air control frequency, and activates homing aids for the return to ship. SCS-81 CIC (Combat Information Center) contacts Yankee Flight and reports that CIC is tracking them on radar. After crossing the point defense zone of the SCS force, non-recoverable munitions are released. Avionics are selectively shut down, vertical take-off/landing mode is selected on the primary flight display, and transition to vertical flight is accomplished. Yankee Leader executes a vertical landing, taxis, and shuts down.

A.2 DECK LAUNCHED INTERCEPT

The four VFA-V/STOL strike aircraft assigned to SCS-81 (Sea Control Ship) are being maintained in a condition 1 alert. Three of the aircraft are configured for immediate commitment on Deck Launched Intercept missions. Two classes of air intercept missiles, two air to air self-defense missiles, and 1000 rounds of 30 mm ammunition comprise the stores load. The modular avionics and software loaded into working memory are both geared to air-to-air engagements. Enemy action is definitely expected, and a general briefing is conducted and updated at every watch change. The U.S. commander expects the enemy to commit its most sophisticated equipment and effective tactics to the attack. The most common enemy tactic in this situation consists of launching long-range ASCM (Anti-Ship Cruise Missiles) from a bomber class aircraft. This aircraft is almost invariably accompanied by a reconnaissance/intelligence aircraft which has no offensive capability to speak of but provides essential targeting information to the launch platform. AEW (Airborne Early Warning) helicopters from the SCS and patrol frigates of the task group have been on station almost continuously for the past two days since the task group is easily within the range of enemy shore-based bombers and fighters.

At 1520 a remote air tracker in the SCS Combat Information Center reports two new contacts entering the northeast quadrant. The air tracker is monitoring returns from an AEW helicopter, and the range of the new contacts is in excess of 200 nmi from task group center. After a single update, the ship's track supervisor assigns the contacts a firm track number and a designation of UNKNOWN AIR. The strength of the returns, coupled with the detection range, indicates the large radar cross sections of bomber aircraft. Considering the criticality of a minimum launch delay for the DLI, the ship's weapon coordinator, CIC supervisor, and OOD (Officer of the Deck) quickly authorize two VFA-V/STOL's to investigate the potential threat. Two pilots, code named Echo 1 and Echo 2, scramble to their planes and run up the engines immediately. No time-consuming equipment initialization is attempted; a single 12-second system GO-NO GO test indicates both aircraft in a GO status. The existence of a condition 1 alert ensures that the deck is clear, and since no appreciable wind is over the deck, the aircraft taxi and execute maximum power short take-offs within 15 seconds of each other. Table A-2 lists the mission phases.

TABLE A-2. DECK LAUNCHED INTERCEPT - MISSION PHASES

1.0 PRE-FLIGHT
2.0 LAUNCH
3.0 CLIMB TO ALTITUDE
4.0 ACCELERATE TO DASH SPEED
5.0 FLY TO ENGAGEMENT AREA
6.0 PRE-ATTACK
7.0 ATTACK (Missile)
8.0 RE-ATTACK (IR Missile)
9.0 BREAK-OFF AND DESCENT
10.0 INBOUND CRUISE
11.0 RECOVER
12.0 POST-FLIGHT

Upon leaving the deck, Echo 1 initiates a constant speed climb to dash altitude. He retracts his gear, and his first cockpit check is the emergency/warning multi-purpose display. This display informs him with an alphanumeric message that one of the back-up navigation components has failed. The failure is evaluated as non-critical since navigation is a minor concern on the air intercept mission where the strike aircraft will be vectored to the engagement area by CIC. Echo 1 clears the message from the display.

Monitoring the VSD (Vertical Situation Display), Echo 1 receives the command to begin leveling off and increasing power for acceleration to dash speed. At this point, Echo 1 is no longer primarily concerned with the vertical situation, and he elects to display an expanded horizontal situation on the upper central CRT. The tactical situation depicted allows Echo 1 to identify all surface traffic, air traffic, and tactically significant points in the vicinity. Echo 1 checks Echo 2's status via UHF: both turn their radars to STANDBY and maintain the assigned bearing to the intercept area.

CIC contacts Echo Flight over the secure voice channel with a tactical update and further instructions. At least four individual tracks comprise the wave approaching the SCS task group. No valid responses have been received to electronic identification attempts, and the tracks have been assigned the HOSTILE-AIR designation. The AEW helicopter will attempt a more positive identification and then fall back toward the ship. The ship's weapons coordinator informs Echo Flight that the enemy group probably consists of two bombers, and two fighter escorts. One bomber will be performing reconnaissance and intelligence gathering while the other will serve as the launch platform for anti-ship cruise missiles. Supporting this expectation is the fact that two of the tracks are widely separated from the other two – a common tactic since the intelligence gathering aircraft has a greater standoff capability than the actual launch platform. It is extremely unlikely that the VFA-V/STOL's can successfully engage the bombers unless the fighter escorts are nullified. For this reason the carrier

task force has dispatched two F-14 CAP aircraft to counter the hostile fighters. The F-14's speed should allow them to engage the opposing fighters before the VFA-V/STOL's arrive in the intercept area.

The basic tasks required of Echo 1 on the cruise out are monitoring flight and subsystem status and preparing for the anticipated intercept. MASTER ARM is actuated. Echo 1 assumes that he will perform a forward hemisphere intercept with the first missile launched just inside the outer boundary of its launch zone. Two semiactive missiles are selected and pre-set. Echo 1 selects an air-to-air mode for his displays and activates his helmet mounted sight in case it may be required later. The radar is switched to AIR SEARCH for two scans, and AUTO TRACK is next depressed to check out vital functions. The radar is then put back in STANDBY, and Echo 1 and Echo 2 maintain heading and maximum speed to the intercept area.

CIC contacts Echo Flight and reports that the tracks have been confirmed as HOSTILE-AIR and almost certainly are the presumed bombers and fighters. Echo 1 is instructed to engage the intelligence aircraft, and Echo 2 is assigned the launch platform. The ship's weapons coordinator also gives Echo Flight confirmation that the F-14's will arrive prior to the VFA-V/STOL's to engage the fighter escorts. If by chance the VFA-V/STOL interceptors are challenged by the enemy fighters, they will attempt to escape and, if necessary, concentrate on intercepting the enemy ASCM's in flight. Echo 1 is given a minor heading correction to the reconnaissance aircraft, and Echo 2 is instructed to remain on his present heading. The ship's weapons coordinator informs the pilots that their tactical displays should soon include the enemy tracks and that he is transferring control of the VFA-V/STOL's to shipboard AIC (Airborne Intercept Controllers).

Echo 1 rolls into his heading change and increases the scale of his tactical presentation. The first two enemy tracks appear on the upper right of his display. Although intercept geometry and tactics are being handled by the AIC, Echo 1 computes alternate solutions on the tactical display for his own benefit. Soon, all four enemy tracks are present on the tactical display and Echo 1 commands his radar to AIR SEARCH. He augments the tactical presentations available to him by presenting the radar video on the VSD. Vertical situation information is still available on the HUD (Head-Up-Display) and the borders of the VSD. Two friendly air symbols move onto the display, and Echo 1 confirms that these tracks are the supporting F-14's. The hostiles must interrupt their escort mission and move out to meet the F-14's as far removed from the bombers as possible. Echo 1 is quickly approaching engagement range.

Although the target is within range as is evident from the tactical display, Echo 1's radar is very noisy and shows no target returns. Checking his threat detection and warning display, Echo 1 realizes that the recon/intell is engaged in active electronic countermeasures. He activates the radar antijamming feature and instructs the computer to derive range data both in its primary mode and synthetically. The radar presentation clears up sufficiently to allow Echo 1 to distinguish a target. He interrogates the target and confirms it as hostile. Decelerating in anticipation of a forward hemisphere attack, Echo 1 verifies that the previously selected missile is the correct choice. Echo 1 designates the track on his radar display and energizes the missile's active tracker. Establishing the proper heading for missile launch, he performs the final settings and releases the missile immediately after receiving the in-range and lock-on indications. Echo 1 quickly activates the second missile's seeker and verifies that the radar target is still designated. He watches his radar for any indication of missile impact, but his threat detection and warning set informs him that the adversary is still jamming. He is now within visual range and the first missile was obviously unsuccessful. Echo 1 releases the second missile and as it approaches the bomber it breaks lock, pitches upward, and harmlessly bypasses the target. Echo 1's last alternative is a rear hemisphere attack. He puts his aircraft into a turn, pulling as many g's as he dares. Exiting the turn he must accelerate and change heading once again to put himself into position for a rear-quarter attack on the fleeing bomber. For this attack he selects an IR passive homing missile even through this particular aircraft presents a less than optimum IR signature. He activates his helmet mounted sight system, tracks the target steadily, receives verification that the missile head has started to track, and launches the missile. The enemy is severely damaged by the impact and is unable to maintain control.

Echo 1 calls Echo 2 on UHF to offer assistance if required. Echo 2 reports that the ASCM launch aircraft has not been destroyed but was forced to break off its mission without launching a missile. Echo 2 also reports that one of the hostile fighters has been destroyed and the other escaped undamaged and may still be in the area. Echo 1 decides to drop to a safer altitude and return to ship. He is provided automatic steering commands by designating the SCS symbol on the tactical display. Both Echo 1 and Echo 2 scan the sky very carefully for any sign of the remaining fighter.

The cruise back is relatively uneventful. Visibility is good, and no enemy interference is experienced. About 100 miles from ship Echo 1 requests CIC to ready a SAR helicopter in case he must eject prior to reaching home. His fuel supply is extremely low, and Echo 1 jettisons all remaining munitions and checks his survival gear. He reviews survival procedures via a call-up on a multipurpose display. About 20 miles out Echo 1 decides he has sufficient fuel to complete the mission. He configures his aircraft, selects take-off/landing mode on the VSD, transitions from conventional flight, and executes a routine short landing.

DISCUSSION

GREEN: (United Kingdom)	Would you defend your statement that VSTOL aircraft are high-workload aircraft? Your presentation seemed to show (figure 7a) that workload peaks did not occur during takeoff and landing, yet when in normal flight VSTOL aircraft behave exactly as any other aircraft.
LINTON: (United States)	I agree that VSTOL aircraft are not inherently higher workload during conventional mission phases. The emphasis in the study was not on takeoff and landing, so you are interpreting figure 7a correctly. Your impression that I reported high-workload in VSTOL was gained from my introductory remarks, which were general and based on more recent findings.
STRASSER: (Germany)	Your model of workload seems to be restricted to time studies and to the concept that time of activation or time to perform a required task is an indicator of workload. Does not this seem to be a possible false conclusion, because working depends on individual strategies of work? Is not the particular "needed time" to perform a required task only one dimension in the complexity of workload?
LINTON: (United States)	While it is true that other factors besides time must be considered, the concept of workload implies "busy-ness" which is time-based. The other dimensions must somehow be reduced to a time base if they are to be relatable to workload as used in our paper.
NICHOLSON: (United Kingdom)	Your model relating performance and workload has certain features which are different from the usual model of Yerkes and Dodson. For instance, there is a wide range of workload with maximum performance. Is this essentially an integration? Relationships between performance--particularly peak performance--and workload may vary for different tasks. This is important, because one could have a peak performance for a specific task which may be the whole purpose of the mission--say, weapons aiming.
LINTON: (United States)	The curve shown was meant to reflect a generalized concept. Whether the center spread of performance stability will be as large as shown remains to be seen, as well as other aspects of the curve as shown. The curve was also meant to reflect multi-task performance, i.e., perhaps an integrated index of all operator behavior within a crew station at some point in the mission.
NICHOLSON: (United Kingdom)	I would agree. We carried out a study on simple reaction time and adaptive tracking during recovery from a sedative drug. There was some evidence that performance recovery with time was different for each task, and that this may be related to optimum arousal for optimum performance on each task.
LINTON: (United States)	It is this type of molecular data on behavioral characteristics which needs to be evaluated for its implications for workload in an integrated multi-task situation found in advanced crew stations.
NICHOLSON: (United Kingdom)	I just want to add that since that study, comparisons have been made with the co-pilot (Nicholson, et al., Aerospace Medicine).

MATHEMATICAL ANALYSIS AND COMPUTER SIMULATION IN MILITARY MISSION WORKLOAD ASSESSMENT

Richard A. Albanese, M.D.
USAF School of Aerospace Medicine
Brooks AFB TX 78235

SUMMARY

This report discusses mathematical and computer approaches to the assessment of crew workload during military missions. Three analysis tools are discussed: (1) estimation algorithms, (2) linear models, (3) nonlinear/hybrid models. These separate but interacting methods provide increasing levels of detail in an analysis, but require increasing levels of effort to obtain a result.

Estimation algorithms are computer programs which provide coarse overall mission analyses. These algorithms use methods from applied probability theory, data from laboratory experiments, and estimates derived by employing subjective estimation techniques such as the Delphi Method.

The linear models use linear ordinary differential equations to represent human and aircraft responses. These methods have been employed to compute statistical properties of flight envelopes and estimates of crew aircraft control effort.

Hybrid models are analysis tools which can treat nonlinear control maneuvers and discrete crew task performance such as decision making, vigilance efforts, and avoidance responses. These models have been used for detailed examination of short critical flight segments. The typical model contains a control loop with pilot equations, aircraft control equations, and aircraft equations of motion. Interacting with the control loop is a task network which represents discrete decision tasks and vigilance tasks as performed by the pilot and/or other crew members depending on the system being studied. These models provide a variety of outputs including actual aircraft trajectories, estimates of critical discrete task performance, and estimates of control effort.

A central element in this presentation is emphasis on the notion that a significant measure of crew workload is acceptable accomplishment of the mission plan. Mathematical analysis and computer simulation, properly employed, provide several measures of workload to the investigator, including the emphasized mission accomplishment measures.

INTRODUCTION

In this report, mathematical and computer approaches to the assessment of crew workload during military missions are described with the focus on methods which attempt to determine whether a typical aircrew can accomplish the workload presented to it by an aircraft weapon system and mission requirements. All numerical values and equation parameters are hypothetical, unless referenced. The equations used are simple, wherever possible, neglecting complications which would interfere with the discussion of the fundamental approach.

ESTIMATION ALGORITHMS

A large portion of the basic data concerning human workload effects consists of laboratory studies of human performance on specific psychomotor tasks. Since the laboratory tasks and environment often differ significantly from the operational tasks and environment of concern, one frequently encounters difficulty when attempting to meaningfully relate laboratory data to the field. Estimation algorithms employ subjective estimation techniques such as the Delphi Method to extrapolate from the laboratory to the field environment (1,2).

Suppose we have gathered data from a particular laboratory psychomotor task and have developed a performance measure of efficiency $E(t)$ which ranges from 0.0 to 1.0 as a function of time. When $E(t) = 1.0$, 100% correct response per unit time is indicated, while $E(t) = 0.0$ indicates no performance on the laboratory task. Our approach has been to fit a dynamic model to data of this sort, whenever possible. For periods of work activity we will use the equation,

$$\frac{dE(t)}{dt} = -\alpha E(t) \tag{1}$$

and for rest or recovery periods we will use

$$\frac{dE(t)}{dt} = r(1.0 - E(t)) \qquad (2)$$

These are first order models which are approximations for the purpose of illustration. Particularly, these equations neglect startup and end-spurt effects (3). The coefficient α is a measure of activity level during a work period, such that $\alpha=0$ indicates no fatiguing and $\alpha > 0$ indicates a fatiguing effect. We will assume $\alpha = 0.00068\ \text{min}^{-1}$ for the laboratory task, corresponding to a 15% decrease in performance efficiency during four hours of continuous performance. The coefficient r is a recovery coefficient. We will assume $r=0.0048\ \text{min}^{-1}$ corresponding to a 90% recovery of performance efficiency after eight hours of rest.

Now our challenge is to project these laboratory data onto actual field conditions. Let us consider a simplified hypothetical strategic bomber mission. The mission can be divided into several segments based on imposed workload. A coarse division could include take-off, climb-out, cruise, aerial refueling, penetration and bomb-drop segments. Two questions are immediate. First, how do the activity levels in these mission segments compare with the activity level of the laboratory task? Second, if during a mission segment, the average performance efficiency of the crew is E_{ave}, as determined on the laboratory psychomotor task, what is the probability of accomplishing this mission segment? A workload analyst can get useful answers to these questions by carefully asking properly prepared subjects to provide subjective estimates. Specifically, the mission segments can first be ranked according to activity level, then numerically scaled relative to the laboratory task. Further, subjects can provide estimates of the probability of proper mission segment completion as a function of segment average performance efficiency, as shown in Figure 1. The curves in Figure 1 are provided by the author for purposes of illustration only.

An analysis of the hypothetical strategic bomber mission is shown in Table I. The mission segment durations and activity levels are needed to calculate average E using equation (1), and were chosen purely for illustration. The average E values were used with the curves of Figure 1 to provide the segment completion probabilities. In this analysis we have assumed continuous activity with no crew rest periods, so that equation 2 was not used. More detailed analyses, adapting the method outlined by Table 1, can provide useful data for mission planning and crew work/rest scheduling for existing systems, and can provide assessments for systems in the design phase.

Tactical air missions can also be approached using estimation techniques. Let us divide a hypothetical tactical mission into seven segments as shown in Table II, where segment durations and activity levels are also found. For this analysis we will not use P versus average E curves as before. Rather, let ρ be the average number of ground weapon systems destroyed by an aircraft per sortie, and let ν be the average number of aircraft lost per sortie due to ground weapon action. We would expect ρ to fall and ν to increase with decreased pilot efficiency during the target engagement segment of the mission. Suppose that

$$\rho = 1.2\ E_{ave} \qquad (3)$$

and

$$\nu = 0.05/E_{ave}^2 \qquad (4)$$

for purposes of illustration, where again, E_{ave} is the average pilot efficiency, relative to our laboratory psychomotor task, during the target engagement segment.

Using equation (1) and equation (2) to calculate E_{ave} for use in equations (3) and (4), we are in position to estimate the impact of fatigue and work-rest cycles on mission outcome. For example, if an airman is required to fly five missions in succession as given in Table II, his average *kill rate (average* ρ) per mission is 1.082, and his average loss rate (average ν) per sortie is 0.062. If four missions are flown in sequence, the average ρ per sortie is 1.098, and the average ν per sortie is 0.060. Data such as these can aid trade-off analyses between crew-manning costs and aircraft costs.

LINEAR MODELS

Manual control of the aircraft is a task which is often important in workload assessment, and there are useful and interesting analytical tools for this purpose. A quasi-linear approach to a manual control task is diagrammed in Figure 2. In this diagram Y_p and Y_c are transfer functions of the human operator and aircraft respectively. The signal $n(t)$ represents signal injected by the human operator into the control loop, but which is not linearly correlated with the system perturbing signal $i(t)$. The signals $s(t)$ and $e(t)$ are important in workload assessments as $s(t)$ *characterizes* the amount of control effort expended by the operator, and $e(t)$ the control he achieves. For $e(t)$ we have

$$f_e(w) = \left|\frac{1}{1 + Y_p Y_c}\right|^2 f_i(w) \tag{5}$$

in the case where $R=0$ and $n(t)=0$. The functions $f_e(w)$ and $f_i(w)$ are the spectral density functions of $e(t)$ and $i(t)$ respectively. These are real-valued functions such that, for example, $f_e(w_2+w_1/2) \cdot (w_2-w_1)$ approximates that portion of the variance in $e(t)$ due to spectral components in $e(t)$ with frequencies between w_1 and w_2 (4). Thus we have

$$\text{VARIANCE OF } e(t) = \int_0^{w_\beta} \left|\frac{1}{1 + Y_p Y_c}\right|^2 f_i(w)\, dw \tag{6}$$

and similarly,

$$\text{VARIANCE OF } s(t) = \int_0^{w_\beta} \left|\frac{Y_p}{1 + Y_p Y_c}\right|^2 f_i(w)\, dw \tag{7}$$

Both equation 6 and 7 assume $R=0$ and $n(t)=0$ and w_β is the bandwidth of the system perturbing signal $i(t)$. If we use the well known cross-over model, we have

$$Y_p Y_c = \frac{w_c e^{-j\tau w}}{jw} \tag{8}$$

where $j = \sqrt{-1}$ and w_c and τ are parameters which depend on Y_c and $w\beta$. Let the transfer function $Y_c = 3/jw$ relate pounds force on the control stick to aircraft pitch angle in degrees. If the perturbing signal $i(t)$ is a narrow bandwidth white noise signal with variance 6 degrees ^{2} and bandwidth 2.5 radians, we have $Y_p = 1.43\ e^{-0.1975jw}$ as the transfer function relating error in degrees to pounds stick force (5,6,7). Substituting into equation 6 and 7 we find that the variance of $e(t)$ is 0.855 degrees^2, while the variance of $s(t)$ is 1.743 pounds^2.

The above example shows how the quasi-linear approach can provide estimates of the degree of control an operator can achieve with a system, and estimates of what physical movement is required. The literature provides a variety of mathematical models for the human operator including more detailed transfer function models and optimal control models (8,9). These models can be used as above, or can be used to provide example flight paths and control movements as functions of time.

HYBRID MODELS

Hybrid models are more complex analysis tools and can treat nonlinear control maneuvers and discrete task performance such as decision making, vigilance and avoidance responses. We have used this model type for detailed simulation and examination of critical flight segments.

To provide an example let us consider again the manual control task treated in the last section, but this time consider that the task is interrupted for λ seconds N times each minute. We will assume the worst case, namely: while performing the secondary task, the operator completely ceases to monitor the loop error signal and stops all control movement, and between performance of the secondary task the operator employs his usual control transfer function. The data in Table III were generated via a simulation model. As expected, the variance of the error signal, VAR(e), increases with increased interruption, N. Note that this increase in variance occurs despite an increase in control effort.

CONCLUSION

Three overlapping classes of analytical tools have been surveyed, viz: estimation algorithms, linear models, and hybrid models. Examples have indicated how these tools can aid in workload analysis. In the examples, the model outputs ranged from probability of mission segment completion to the variance of aircraft pitch angle control. These are mission relevant workload metrics, and as such are meaningful in an operational context. At the very least, mathematical methods enable a workload analyst to provide internally consistent workload assessments to using elements. In the best circumstance, mathematical methods can effectively bridge the gap between the laboratory and field needs.

REFERENCES

1. Dalkey, N.C. The delphi method: an experimental study of group opinion. RM-5888-PR, Rand, Santa Monica, CA, 1969.

2. Albanese, R. and Pickering, J.E. Aircrew vulnerability in nuclear encounters. Military medicine. Vol. 139, No. 12, Dec. 1974, p 945-951.

3. Shephard, R.J. Men at work. Applications of ergonomics to performance and design. Illinois: Charles C Thomas, 1974, p 217-218.

4. Albanese, R.A. Reliability of human control. In F. Proschan and R.J. Serfling, eds. Reliability and Biometry, Statistical Analysis of Lifelength, SIAM, Philadelphia, 1974, p 727-742.

5. McRuer, D.T. and E. Krendel. Human pilot dynamics in compensatory systems. AFFDL-TR-65-15, Jul. 1965.

6. McRuer, D.T. and H.R. Jex. A review of quasi-linear pilot models. IEEE Trans., Vol HFE-8, Number 3, Sept. 1967, p 231-249.

7. McRuer, D.T. and E.S. Krendel. Mathematical models of human pilot behavior. AGARD-AG-188, Jan. 1974.

8. Kleinman, D.L., Baron, S., and Levison, W.H. A control theoretic approach to manned-vehicle systems analysis. IEEE Trans. Automatic Control, AC-16, 1971, p824-832.

9. Sheridan T.B., Ferrell, W.R. Man-machine systems: Information, control, and decision models of human performance. Massachusetts: MIT Press, 1974.

TABLE I

Mission Segment	Segment Duration (minutes)	Segment Activity Level (min^{-1})	Average E during Segment	Probability of Proper Segment Completion
Take-off	5	0.00136	0.997	0.9994
Climb-out	40	0.00068	0.980	0.9980
Cruise	180	0.00006	0.961	1.0000
Refuel	30	0.00136	0.937	0.9874
Penetrate	120	0.00102	0.864	0.9864
Bomb-drop	20	0.00102	0.803	0.9803

TABLE II

Mission Segment	Segment Duration (minutes)	Segment Activity Level (min^{-1})
Pre-mission Planning	20	0.00014
Take-off	5	0.00136
Fly-to Target	40	0.00014
Engage Target	15	0.00170
Return Cruise	50	0.00014
Land	20	0.00102
Between Mission Rest	60	------

TABLE III

λ(seconds)	N	VAR(e) ($degrees^2$)	VAR(s) ($pounds^2$)
0.3	0	0.855	1.743
0.3	12	0.985	1.995
0.3	30	1.185	2.345
0.3	60	1.560	3.535

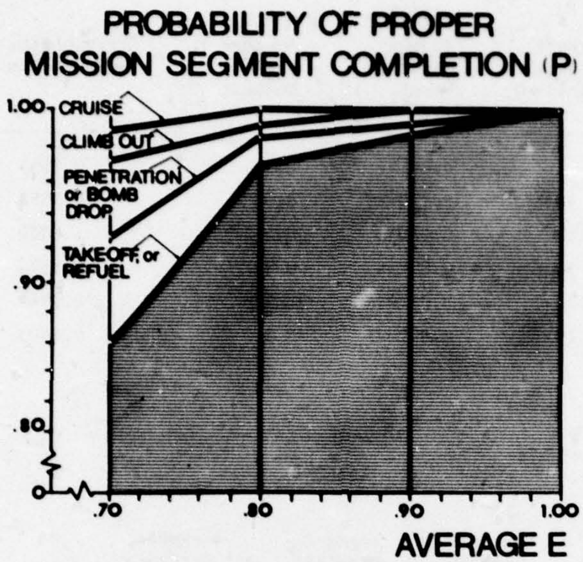


FIGURE 1: P versus average E for a hypothetical bomber mission.

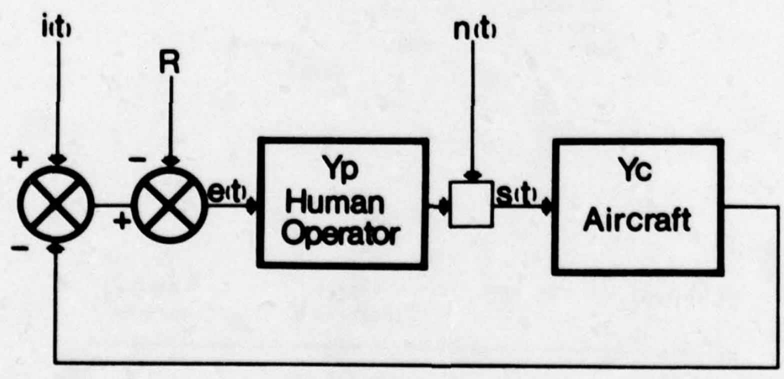


FIGURE 2: A quasi-linear approach to a manual control task.

DISCUSSION

JAHNS: (United States)
In your "basic concept" (last vu-foil), mission success being an indicator of crew workload does not appear to follow traditional definitions of workload. It seems to be more related to performance capability and limitations of the crew. Would you clarify, please?

ALBANESE: (United States)
I am suggesting that mission success rate is a measure of workload. Certainly if a mission fails because the crew could not accomplish the required activities, this is a workload problem.

NICHOLSON: (United Kingdom)
How did you reach attrition rates for current tactical air missions? And, how would you predict attrition rates for interdiction-strike missions?

ALBANESE: (United States)
The attrition rate versus efficiency curves shown are hypothetical curves for illustrative purposes only. Obtaining these curves is very difficult. Possible sources are: study of mission records, detailed simulations, operations analysis and subjective estimates.

KIMBALL: (United States)
As a followon to Wing Commander Nicholson's question, what level of success or reliability have you empirically determined over the profiles to which you have applied these models, say for example in simulated inflight sorties or a number of simulated combat encounters?

ALBANESE: (United States)
The bomber mission estimation algorithm has had limited validation in the laboratory setting. The hybrid model, applied to aerial refueling, has had some validation against field data.

ROUND TABLE DISCUSSION

Technical difficulties in the recording system prevent a complete account of the round table discussion. What follows is the symposium chairman's summary based on recall, and probably containing some elements of discussion after papers. This is not totally inappropriate, since the level of audience participation in both workload symposia was high and of considerable technical merit. Five questions were addressed to the panel. The panel membership consisted of the speakers.

Question #1. Test devices, methods, and statistical evaluation techniques vary from study to study. What is the common ground and how do we achieve it so that we can reach generalizations regarding pilot workload?

This is a challenging issue for most major problem areas in aerospace medicine and its related disciplines in biotechnology. The common ground, of course, is that the basic model of man has not changed. Also, we should assume that the more general the generalizations are, the greater the likelihood that we should see congruence at the level of findings from different approaches. However, commonality is frequently not achieved because the research questions being addressed lead to significantly different endpoints. Part of our job as advisors to the operational community is to bring bits and pieces together in a framework of logic which is defensible and to make the generalizations for our audience. This is difficult. It is far easier to stay as close as possible to the operational environment in our research; this minimizes the problem. (Editor's note: Consider this last point in relation to the TER at the beginning of this document.)

Question #2. What are the relationships between the psychomotor, the psychometric, the psychophysiologic, and the physiological methods and measurements? Do they lead to common findings? Qualitative or quantitative?

It seems unlikely that we will ever bring these domains together. Indeed, perhaps, we shouldn't try. Each has it advantages and limitations. A careful assessment of the kind of workload question being asked and the constraints existing in the most appropriate or available research environments (in the cockpit vs. in the laboratory) leads the experimenter to a choice. Forcing all approaches through a final "funnel" to yield a small number of "truths" about workload is probably counter-productive. The larger goal is to provide the best possible response to a request from our operational customers.

Question #3. How does one choose, relate, and apply subjective versus objective methods?

It would be fun to debate the theoretical issues--for example, the virtues of the Cooper-Harper rating scale on aircraft handling qualities vs. the HOS described in one of the papers in the symposium--but our experience is that the choice is based on experimenter skills and resources and constraints in the environment in which the experiment is to be conducted. All else follows this choice. However, the panel endorses the trends toward greater use of objective measurement in the most realistic environment--in the cockpit--for many workload questions. The several papers on helicopter aircrew performance presented by the Ft. Rucker group exemplify this quite well.

Question #4. How do conventional workload measures relate to measures and methods derived from mathematical modelling approaches?

This question can't be answered unless there is a prior determination that the two groups are addressing similar questions. Given that, the problem becomes one of interaction and communication, which has not occurred sufficiently in the past. There are key issues to be addressed. The experimenter must become more active in feeding experimental data to the modeller and the modeller must be more informative as to his needs. It is encouraging to see more interaction and communication, to see the experimenter exploiting the inherent advantages of the model, which can lead to significant findings with minimal data or missing data. But as the last paper today demonstrated, there are times when we must conclude that the final result will not infrequently be a set of metrics new to the experimenter. His task is to learn to understand and use them without prejudice.

Question #5. What workload findings should be delivered to the operational community, the design community, the human engineering (human factors/ergonomics) community? Are our data of any use to anyone except ourselves? When we are asked for a "position," is it based on our opinion, our interpretation of our data, or data offered for interpretation by anyone?

Obviously, the chairman intended this question to be provocative--perhaps even provoking. There is a temptation to ignore it, or at least parts of it. But there is an issue of merit: most of us here represent laboratories whose funds come from some national defense budget. What is that nation's "return on investment?" It is clear from the papers of the past day and a half that we are addressing substantive questions, and that our products should go to several different customers: the operators, the designs, the commanders, to name a few. It is also clear that a substantial part of our work is in response to a request. A potential dilemma exists for the rest of the work--if our work is self-initiated, addressing a need we perceive even if others don't, then it can turn out that no one wants the answers. Then the "return on investment" drops to a very low level.

A-15-1

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-CP-216	ISBN 92-835-1285-5	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title

METHODS TO ASSESS WORKLOAD

7. Presented at the Aerospace Medical Panel Specialists' Meeting held in Cologne, FRG, 18–22 April 1977.

8. Author(s)	9. Date
Various	June 1978

10. Author's Address	11. Pages
Various	142

12. Distribution Statement This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors

Work measurement	Flight crews	Man machines systems
Operators (personnel)	Performance evaluation	Meetings
Pilots (personnel)	Human factors engineering	

14. Abstract

For the measurement of workload a wide variety of methods is employed, frequently in an interdisciplinary setting.

This multifaceted approach to the problems of the methods best suited to assess workload is highlighted by the range of topics covered in the meeting.

Methodology questions involve basic issues like measurement sensitivity, reliability and validity; instrumentation and associated techniques; study environment ranging from the laboratory to simulator and in-flight studies.

AGARD Conference Proceedings No.216
Advisory Group for Aerospace Research and Development, NATO
METHODS TO ASSESS WORKLOAD
Published June 1978
142 pages

For the measurement of workload a wide variety of methods is employed, frequently in an interdisciplinary setting.

This multifaceted approach to the problems of the methods best suited to assess workload is highlighted by the range of topics covered in the meeting.

Methodology questions involve basic issues like measurement sensitivity, reliability and validity; instrumentation

P.T.O.

AGARD-CP-216

Work measurement
Operators (personnel)
Pilots (personnel)
Flight crews
Performance evaluation
Human factors engineering
Man machines systems
Meetings

AGARD Conference Proceedings No.216
Advisory Group for Aerospace Research and Development, NATO
METHODS TO ASSESS WORKLOAD
Published June 1978
142 pages

For the measurement of workload a wide variety of methods is employed, frequently in an interdisciplinary setting.

This multifaceted approach to the problems of the methods best suited to assess workload is highlighted by the range of topics covered in the meeting.

Methodology questions involve basic issues like measurement sensitivity, reliability and validity; instrumentation

P.T.O.

AGARD-CP-216

Work measurement
Operators (personnel)
Pilots (personnel)
Flight crews
Performance evaluation
Human factors engineering
Man machines systems
Meetings

AGARD Conference Proceedings No.216
Advisory Group for Aerospace Research and Development, NATO
METHODS TO ASSESS WORKLOAD
Published June 1978
142 pages

For the measurement of workload a wide variety of methods is employed, frequently in an interdisciplinary setting.

This multifaceted approach to the problems of the methods best suited to assess workload is highlighted by the range of topics covered in the meeting.

Methodology questions involve basic issues like measurement sensitivity, reliability and validity; instrumentation

P.T.O.

AGARD-CP-216

Work measurement
Operators (personnel)
Pilots (personnel)
Flight crews
Performance evaluation
Human factors engineering
Man machines systems
Meetings

AGARD Conference Proceedings No.216
Advisory Group for Aerospace Research and Development, NATO
METHODS TO ASSESS WORKLOAD
Published June 1978
142 pages

For the measurement of workload a wide variety of methods is employed, frequently in an interdisciplinary setting.

This multifaceted approach to the problems of the methods best suited to assess workload is highlighted by the range of topics covered in the meeting.

Methodology questions involve basic issues like measurement sensitivity, reliability and validity; instrumentation

P.T.O.

AGARD-CP-216

Work measurement
Operators (personnel)
Pilots (personnel)
Flight crews
Performance evaluation
Human factors engineering
Man machines systems
Meetings

and associated techniques; study environment ranging from the laboratory to simulator and in-flight studies.

The papers were presented at the Aerospace Medical Panel Specialists' Meeting held in Cologne, FRG, 18–22 April 1977.

ISBN 92-835-1285-5

and associated techniques; study environment ranging from the laboratory to *simulator and in-flight studies.*

The papers were presented at the Aerospace Medical Panel Specialists' Meeting held in Cologne, FRG, 18–22 April 1977.

ISBN 92-835-1285-5

and associated techniques; study environment ranging from the laboratory to simulator and *in-flight* studies.

The papers were presented at the Aerospace Medical Panel Specialists' Meeting held in Cologne, FRG, 18–22 April 1977.

ISBN 92-835-1285-5

and associated techniques; study environment ranging from the laboratory to simulator and in-flight studies.

The papers were presented at the Aerospace Medical Panel Specialists' Meeting held in Cologne, FRG, 18–22 April 1977.

ISBN 92-835-1285-5